George A. Tsihrintzis
Lakhmi C. Jain
Editors



Multimedia Services in Intelligent Environments

Integrated Systems

SMART INNOVATION, SYSTEMS AND TECHNOLOGIES 3

Springer

George A. Tsihrintzis and Lakhmi C. Jain (Eds.)

Multimedia Services in Intelligent Environments – Integrated Systems

Smart Innovation, Systems and Technologies 3

Editors-in-Chief

Prof. Robert James Howlett
KES International
PO Box 2115
Shoreham-by-sea
BN43 9AF
UK
E-mail: rjhowlett@kesinternational.org

Prof. Lakhmi C. Jain
School of Electrical and Information Engineering
University of South Australia
Adelaide, Mawson Lakes Campus
South Australia SA 5095
Australia
E-mail: Lakhmi.jain@unisa.edu.au

Further volumes of this series can be found on our homepage: springer.com

Vol. 1. Toyoaki Nishida, Lakhmi C. Jain, and Colette Faucher (Eds.)
Modeling Machine Emotions for Realizing Intelligence, 2010
ISBN 978-3-642-12603-1

Vol. 2. George A. Tsihrintzis, Maria Virvou, and Lakhmi C. Jain (Eds.)
Multimedia Services in Intelligent Environments –
Software Development Challenges and Solutions, 2010
ISBN 978-3-642-13354-1

Vol. 3. George A. Tsihrintzis and Lakhmi C. Jain (Eds.)
Multimedia Services in Intelligent Environments –
Integrated Systems, 2010
ISBN 978-3-642-13395-4

George A. Tsihrintzis and Lakhmi C. Jain (Eds.)

Multimedia Services in Intelligent Environments

Integrated Systems

Springer

Prof. George Tsihrintzis
Department of Informatics
University of Piraeus
80 Karaoli-Dimitriou Str.
18534 Piraeus
Greece
E-mail: geoatsi@unipi.gr

Prof. Dr. Lakhmi C. Jain
School of Electrical and
Information Engineering
University of South Australia
KES Centre
Mawson Lakes Campus
Adelaide SA 5095
Australia
E-mail: Lakhmi.jain@unisa.edu.au

ISBN 978-3-642-13395-4 e-ISBN 978-3-642-13396-1

DOI 10.1007/978-3-642-13396-1

Smart Innovation, Systems and Technologies ISSN 2190-3018

Library of Congress Control Number: 2010928211

Typesetting: Scientific Publishing Services Pvt. Ltd., Chennai, India.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Foreword

KES International (KES) is a worldwide organisation that provides a professional community and association for researchers, originally in the discipline of Knowledge Based and Intelligent Engineering Systems, but now extending into other related areas. Through this, KES provides its members with opportunities for publication and beneficial interaction.

The focus of KES is research and technology transfer in the area of Intelligent Systems, i.e. computer-based software systems that operate in a manner analogous to the human brain, in order to perform advanced tasks. Recently KES has started to extend its area of interest to encompass the contribution that intelligent systems can make to sustainability and renewable energy, and also the knowledge transfer, innovation and enterprise agenda.

Involving several thousand researchers, managers and engineers drawn from universities and companies world-wide, KES is in an excellent position to facilitate international research co-operation and generate synergy in the area of artificial intelligence applied to real-world 'Smart' systems and the underlying related theory.

The KES annual conference covers a broad spectrum of intelligent systems topics and attracts several hundred delegates from a range of countries round the world. KES also organises symposia on specific technical topics, for example, Agent and Multi Agent Systems, Intelligent Decision Technologies, Intelligent Interactive Multimedia Systems and Services, Sustainability in Energy and Buildings and Innovations through Knowledge Transfer. KES is responsible for two peer-reviewed journals, the International Journal of Knowledge based and Intelligent Engineering Systems, and Intelligent Decision Technologies: an International Journal.

KES supports a number of book series in partnership with major scientific publishers.

Published by Springer, 'Smart Innovative Systems and Technologies' is the KES flagship book series. The aim of the series is to make available a platform for the publication of books (in both hard copy and electronic form) on all aspects of single and multi-disciplinary research involving smart innovative systems and technologies, in order to make the latest results available in a readily-accessible form.

The series covers systems that employ knowledge and intelligence in a broad sense. Its focus is systems having embedded knowledge and intelligence, which may be applied to the solution of world industrial, economic and environmental problems and the knowledge-transfer methodologies employed to make this happen effectively. The combination of intelligent systems tools and a broad range of applications introduces a need for a synergy of scientific and technological disciplines.

Examples of applicable areas to be covered by the series include intelligent decision support, smart robotics and mechatronics, knowledge engineering, intelligent multi-media, intelligent product design, intelligent medical systems, smart industrial products, smart alternative energy systems, and underpinning areas such as smart systems theory and practice, knowledge transfer, innovation and enterprise.

The series includes conference proceedings, edited collections, monographs, handbooks, reference books, and other relevant types of book in areas of science and technology where smart systems and technologies can offer innovative solutions.

High quality is an essential feature for all book proposals accepted for the series. It is expected that editors of all accepted volumes take responsibility for ensuring that contributions are subjected to an appropriate level of reviewing process and adhere to KES quality principles.

Professor Robert J. Howlett
Executive Chair, KES International
Visiting Professor, Enterprise: Bournemouth University
United Kingdom

Editors

George A. Tsihrintzis received the Diploma of Electrical Engineer from the National Technical University of Athens, Greece (with honors) and the M.Sc. and Ph.D. degrees in Electrical Engineering from Northeastern University, Boston, Massachusetts, USA. He is currently a Professor in the Department of Informatics, The University of Piraeus, Greece. His current research interests include Pattern Recognition, Decision Theory, and Statistical Signal Processing and their applications in Multimedia Services, User Modeling, Intelligent Software Systems, Human-Computer Interaction and Information Retrieval. He has authored or co-authored over 200 research articles in these areas, which have appeared in international journals, book chapters, and conference proceedings, and has served as the principal investigator or co-investigator in several R&D projects. He is the sole author of a book on *Image Analysis* (in Greek) and co-author of a book on *Principles and Applications of Signals and Systems* (in Greek) and an upcoming book on *Visual Affect Recognition* (IOS Press, 2010). He has served as a member of Program Committees and/or reviewer of International journals and conferences. He was the founding general and program co-chair of the **2008 International Symposium on Intelligent Interactive Multimedia Systems and Services (KES-IIMSS 2008)**, Piraeus, Greece, July 9-11, 2008, organized jointly by the Department of Informatics of the University of Piraeus and KES International. He was the honorary co-chair of the **2009 International Symposium on Intelligent Interactive Multimedia Systems and Services (KES-IIMSS 2009)**, Venice, Italy, July 16-17, 2009. He is the general and program co-chair of the **2010 International Joint Conference on e-Business and Telecommunications (ICETE 2010)**, Piraeus, Greece, July 26-28, 2010. He is the founding general co-chair of the **2010 International Multi-Conference on Innovative Developments in Information Communication Technologies (INNOV 2010)**, Piraeus, Greece, July 29-31, 2010. He is the general and program co-chair of the **2010 International Symposium on Intelligent Interactive Multimedia Systems and Services (KES-IIMSS 2010)**, Baltimore, USA, July 28-30, 2010. He has co-edited the following books: (1) *Multimedia Services in Intelligent Environments - Advanced Tools and Methodologies*, volume 120 in Studies in Computational Intelligence (SCI) Book Series, Springer 2008, (2) *Computational Intelligence Paradigms - Innovative Applications*, volume 137 in Studies in Computational Intelligence (SCI) Book Series, Springer 2008, (3) *New Directions in*

Intelligent Interactive Multimedia, volume 142 in Studies in Computational Intelligence (SCI) Book Series, Springer 2008, (4) *Multimedia Services in Intelligent Environments - Software Development Challenges and Solutions*, in Smart Innovation, Systems and Technologies (SIST) Book Series, Springer 2010, and (5) *Multimedia Services in Intelligent Environments - Integrated Systems*, in Smart Innovation, Systems and Technologies (SIST) Book Series, Springer 2010. He was a guest co-editor of the special issue on "Intelligent Modelling and Data Analysis Techniques" of the **International Journal of Intelligent Defence Support Systems** (Inderscience, 2009). He was a guest co-editor of the special issues on "Knowledge-based Modes of Human-Computer Interaction" and "Knowledge-based Environments and Services in Human-Computer Interaction" of the **Intelligent Decision Technologies Journal** (IOS Press, 2010). He won the Best Poster Paper Award of the 5th International Conference on Information Technology: New Generations, Las Vegas, USA, April 7-9, 2008, for co-authoring a paper titled: "Evaluation of a Middleware System for Accessing Digital Music Libraries in Mobile Services." He also won one of the Best Applications Papers Award of the 29th Annual International Conference of the British Computer Society Specialist Group on Artificial Intelligence, Cambridge, UK, December 15-17, 2009, for co-authoring a paper titled: "On Assisting a Visual-Facial Affect Recognition System with Keyboard-Stroke Pattern Information." He can be reached at geoatsi@unipi.gr.

Professor Lakhmi C. Jain is a Director/Founder of the Knowledge-Based Intelligent Engineering Systems (KES) Centre, located in the University of South Australia. He is a fellow of the Institution of Engineers Australia.

His interests focus on the artificial intelligence paradigms and their applications in complex systems, art-science fusion, virtual systems, e-education, e-healthcare, unmanned air vehicles and intelligent agents.

Preface

Multimedia services are widely used by humans in their daily lives. Society is increasingly demanding convenient access to the wealth of information that is available, in a friendly environment.

This book is a continuation of previous volumes providing various perspectives on multimedia services in intelligent environments [1-4].

It includes fourteen chapters on integrated multimedia systems and services covering various aspects such as geographical information systems, recommenders, interactive entertainment, e-learning, medical diagnosis, telemonitoring, attention management, e-welfare and brain-computer interfaces

Each chapter in the book was reviewed by two independent reviewers for quality, novelty and clarity of the research presented in it. We are grateful to the authors and the reviewers for their excellent contributions and visionary ideas. This research book is directed to professors, researchers, application engineers and students of all disciplines.

We wish to express our appreciation to the KES Community for supporting KES International and taking part in the events related to their areas of interests. Thanks are due to Springer-Verlag for their excellent support to KES international.

George A. Tsihrintzis, Greece
Lakhmi C. Jain, Australia

References

[1] Tsihrintzis, G. and Jain, L.C. (Editors), Multimedia Services in Intelligent Environments - Advanced Tools and Methodologies, Springer-Verlag, 2008.
[2] Tsihrintzis, G., Virvou, M., Howlett, R.J. and Jain, L.C. (Editors), New Directions in Intelligent Interactive Multimedia, Springer-Verlag, 2008.
[3] Damiani, E., Jeong, J., Howlett, R.J. and Jain, L.C. (Editors), New Directions in Intelligent Interactive Multimedia Systems and Services - 2, Springer-Verlag, Germany, 2009.
[4] Tsihrintzis, G., Virvou, M., and Jain, L.C. (Editors), Multimedia Services in Intelligent Environments – Software Development Challenges and Solutions, Springer-Verlag, 2010.

Table of Contents

Chapter 1

Advances in Multimedia Services in Intelligent Environments – Integrated Systems

George A. Tsihrintzis[1] and Lakhmi C. Jain[2]

[1] University of Piraeus, Greece
[2] University of South Australia
South Australia
Australia

1 Introduction

Multimedia services is the term chosen to describe services which rely on the coordinated and secure storage, processing, transmission, and retrieval of information which exists in various forms [1]. The term refers to the several levels of data processing. It includes application areas, such as digital libraries, e-learning, e-government, e-commerce, e-entertainment, e-health, and e-legal services. In our earlier book [2], we covered aspects of low level data processing of multimedia services in intelligent environments, including storage, recognition, classification, transmission, retrieval, and securing of information. Four additional chapters in [2] considered systems developed to support intermediate level multimedia processing services. These included noise and hearing monitoring, and measurement, augmented reality, and automated lecture rooms. In addition rights management and licensing were included. The final chapter in [2] was devoted to a high-level intelligent recommender service in scientific digital libraries.

The book at hand presents various integrated systems that were developed to accommodate multimedia services in intelligent environments. In addition to this chapter, the present book includes thirteen additional chapters. Chapters 2 and 3 are devoted to multimedia geographical information systems. Specifically, Chapter 2 by Gemizi, Tsihrintzis, and Petalas is on "Use of GIS and Multi-Criteria Evaluation techniques in environmental problems," while Chapter 3 by Charou, Kabassi, Martinis, and Stefouli is on "Integrating multimedia GIS Technologies in a Recommendation System for Geotourism."

Chapters 4 and 5 cover aspects of e-entertainement systems. Specifically, Chapter by El-Nasr and Zupko is on "Lighting Design Tools for Interactive Entertainment," while Chapter 5 by Szczuko and Kostek is on "Utilization of Fuzzy Rules in Computer Character Animation".

Chapters 6 and 7 cover aspects of education and e-learning systems. Specifically, Chapter 6 by Nakatani, Tsumaki, and Tamai is on "Instructional Design of a Requirements Engineering Education for Professional Engineers," while Chapter 7 by Burdescu and Mihăescu is on "Building Intelligent e-Learning Systems by Activity Monitoring and Analysis."

G.A. Tsihrintzis and L.C. Jain (Eds.): Multimedia Services in Intel. Environ., SIST 3, pp. 1–3.
springerlink.com

Chapters 8 and 9 are devoted to medical diagnosis systems. Specifically, Chapter 8 by Schmidt and Vorobieva is on "Supporting the Search for Explanations of Medical Exceptions," while Chapter 9 by Aupet, Garcia, Guyennet, Lapayre, and Martins is on "Security in Medical Telediagnosis."

Chapters 10 and 11 are devoted to telemonitoring systems. Specifically, Chapter 10 by Żwan, Sobala, Szczuko, and Czyzewski is on "Audio Content Analysis in the Urban Area Telemonitoring System," whileChapter eleven by Dalka, Szwoch, Szczuko, and Czyzewski is on "Video Content Analysis in the Urban Area Telemonitoring System."

The next chapter, namely Chapter 12 by Karapiperis, Stojanovic, Anicic, Apostolou and Despotis, is on "Enterprise Attention Management."

An additional chapter, namely Chapter 13 by Raij and Lehto, is on "e-Welfare as a Client-driven Service Concept."

Finally, Chapter 14 by Panoulas, Hadjileondiadis, and Panas is on "Brain-Computer Interface (BCI): Types, Processing Perspectives and Applications."

2 Conclusions

As multimedia services for intelligent environments become increasingly more demanding, new challenges appear which require even more sophisticated tools, methodologies, and integrated systems to solve them. Coincidently, the application areas of multimedia services continue to expand at a very high rate. As a result, the entire field of multimedia services in intelligent environments cannot be effectively covered in one or two volumes. It may be expected that future volumes on other aspects of multimedia services in intelligent environments will continue to appear.

References and Further Reading

[1] Tsihrintzis, G., Virvou, M., Jain, L.C. (eds.): Multimedia Services in Intelligent Environments – Software Development Challenges and Solutions. Springer, Heidelberg (2010)

[2] Tsihrintzis, G., Jain, L.C. (eds.): Multimedia Services in Intelligent Environments – Advanced Tools and Methodologies. Springer, Heidelberg (2008)

[3] Tsihrintzis, G., Virvou, M., Howlett, R.J., Jain, L.C. (eds.): New Directions in Intelligent Interactive Multimedia. Springer, Heidelberg (2008)

[4] Damiani, E., Jeong, J., Howlett, R.J., Jain, L.C. (eds.): New Directions in Intelligent Interactive Multimedia Systems and Services, vol. 2. Springer, Germany (2009)

[5] Wang, F.H., Pan, J.S., Jain, L.C.: Innovations in Digital Watermarking Techniques. Springer, Germany (2009)

[6] Pan, J.S., Chen, Y.W., Jain, L.C.: Intelligent Information Hiding and Multimedia Signal Processing. IEEE Computer Society Press, USA (2009)

[7] Virvou, M., Jain, L.C. (eds.): Intelligent Interactive Systems in Knowledge-Based Environments. Springer, Heidelberg (2008)

[8] Pan, J.S., Niu, X.M., Huang, H.C., Jain, L.C. (eds.): Intelligent Information Hiding and Multimedia Signal Processing. IEEE Computer Society Press, USA (2008)

[9] Camastra, F., Vinciarelli, A.: Machine Learning for audio, Image, and Video Analysis. Springer, London (2008)

[10] Sommerer, C., Jain, L.C., Mignonneau, L. (eds.): The Art and Science of Interface and Interaction Design, vol. 1. Springer, Heidelberg (2008)
[11] Pierre, S. (ed.): E-Learning Networked Environments and Architectures. Springer, London (2007)
[12] Pan, J.-S., Huang, H.-C., Jain, L.C., Fang, W.-C. (eds.): Intelligent Multimedia Data Hiding. Springer, Heidelberg (2007)
[13] Liao, B.-H., Pan, J.-S., Jain, L.C., Liao, M., Noda, H., Ho, A.T.S.: Intelligent Information Hiding and Multimedia Signal Processing, vol. 1. IEEE Computer Society Press, USA (2007)
[14] Liao, B.-H., Pan, J.-S., Jain, L.C., Liao, M., Noda, H., Ho, A.T.S.: Intelligent Information Hiding and Multimedia Signal Processing, vol. 2. IEEE Computer Society Press, USA (2007)
[15] Pan, J.-S., Huang, H.C., Jain, L.C. (eds.): Intelligent Watermarking Techniques. World Scientific Publishing Company, Singapore (2004)
[16] Howlett, R., Ichalkaranje, N., Jain, L.C., Tonfoni, G. (eds.): Internet-Based Intelligent Information Processing. World Scientific Publishing Company, Singapore (2002)

Chapter 2

Use of GIS and Multi-Criteria Evaluation Techniques in Environmental Problems

Alexandra Gemitzi, Vassilios A. Tsihrintzis, and Christos Petalas

Laboratory of Ecological Engineering and Technology,
Department of Environmental Engineering, School of Engineering,
Democritus University of Thrace,
67100 Xanthi, Greece
tsihrin@otenet.gr

Abstract. The application of Multi-Criteria Evaluation in combination with Geographical Information Systems in handling problems related to the environment and to spatial distribution of variables is presented through three case studies. The first one deals with the application of Boolean evaluation in siting natural wastewater treatment facilities, the second one applies both exclusionary and non-exclusionary criteria in siting a Municipal Solid Waste landfill, while the third one computes groundwater vulnerability to pollution applying non-exclusionary criteria coupled with weighted linear combination and ordered weighted averaging. Concluding, the present work highlights the advantages and the limitations of those applications in problems of siting human activities, urban development, resource allocation and management.

Keywords: Multi-Criteria Evaluation, Decision Making Techniques, Geographical Information Systems.

1 Introduction

In order to reach a decision related to a specific objective, several criteria should be evaluated. A scientist usually seeks the best alternative in a multivariate problem, which in most cases is a compromise among various, often controversial, criteria. This procedure is known as Multi-Criteria Evaluation (MCE) [1], [2]. Recently, there is a growing interest on the use of Geographic Information Systems (GIS) as a decision support tool in MCE.

Two are the main procedures most commonly used in achieving Multi-Criteria Evaluation. The first involves Boolean overlay, where all criteria are reduced to logical statements of suitability, i.e., suitable or not suitable, and then combined by means of one or more logical operators such as the AND, known as intersection operator, and the OR, known as union operator. The second procedure is known as Weighted Linear Combination (WLC), where continuous criteria, known as factors or non-exclusionary criteria, are standardized to a common numeric range and afterwards are combined using a weighted average. The result is a continuous suitability map, which may be masked by one or more Boolean constraints and finally thresholded to offer a final decision [3].

G.A. Tsihrintzis and L.C. Jain (Eds.): Multimedia Services in Intel. Environ., SIST 3, pp. 5–62.
springerlink.com

While being well coupled with GIS, these two procedures frequently lead to different results. The Boolean evaluation is an extreme form of decision-making. In the case where criteria are combined with an AND operator, a location must meet every criterion in order to be included in the decision set. If even a single criterion fails to be met, then the location is excluded. In this case the criteria are known as exclusionary criteria or constraints. Such an approach is risk-averse, as it selects locations based on the most cautious strategy possible. It is applied in environmental sciences among others, as it is fast and can easily be applied to regional scale cases, offering, however, a very conservative alternative. On the other hand, if the OR operator is used, a location is included in the decision set even if only a single criterion passes the test. This is a very risk-taking strategy, which is not recommended in environmental decision making.

In the following paragraphs the examination of three environmental problems is presented. The first one deals with the detection of the suitable sites for the construction of natural wastewater treatment facilities, such as stabilization ponds, applying a simple Boolean evaluation. The second one deals with the construction of a suitability map for siting a Municipal Solid Waste landfill. The third one presents the computation of groundwater vulnerability to pollution, examining various environmental and socioeconomic criteria. Both the last two cases apply the Weighted Linear Combination technique.

2 The Boolean Evaluation — Use of GIS in Siting Stabilization Pond Facilities for Domestic Wastewater Treatment

In this section the use of GIS is presented in the problem of siting areas for construction of natural systems such as stabilization ponds (SP) for domestic wastewater treatment. The GIS technology was used in order to create and analyze several grids of different themes and then highlight the areas of interest, i.e., those that satisfy all the relevant criteria. The sizing and performance method calculates the area requirements based on BOD_5 and total coliform effluent discharge criteria as a function of temperature and population. The combination of the above two methodologies, at the municipal level, shows that it is quite simple to check whether natural treatment systems, such as stabilization ponds, can be an alternative wastewater treatment option for each one of the examined municipalities and to calculate the maximum area required for such facilities.

For this purpose, several variables, such as topography, land use, type of geological formation, distance to major rivers or lakes, distance to existing cities and villages, existence of environmentally protected areas, mean minimum monthly temperatures and required wastewater effluent characteristics were analyzed with the GIS using Boolean evaluation, in order to accept or reject a particular area within a region. The method was applied in the region of Thrace (Northeast Greece), an area of approximately 8500 km^2 (Fig. 1). The required area for stabilization pond systems was calculated in each of the 36 municipalities of Thrace (including two islands, Thassos and Samothraki) as a function of the population of each municipality, temperature and local wastewater effluent discharge criteria. Based on the GIS analysis, suitable locations were identified in each municipality first, and then the total required surface area of these systems was compared to the available surface area of each municipality, in order to decide whether stabilization pond systems could be a viable solution to the wastewater management problem in the particular region. In order to cover the entire study area (master plan level), it was necessary to work at a small scale (1:200 000 – 1:250 000, except of the land use data available only at scale of 1:500 000). A more accurate study can be done applying the same methodology at the selected areas using finer scale maps and applying non-exclusionary criteria. Thus, the present methodology offers a fast and simple method to check the suitability of new areas for construction of such systems.

Initially, the process involved the creation and analysis of several grids of different themes [4], [5]. The GIS based methodology resulted in various thematic maps, which satisfy specific suitability, environmental and socio-economic criteria. For each criterion, there are two major discrete categories, i.e., those that include all areas suitable for natural treatment system siting and the ones unsuitable for the specified purpose, i.e., the process applies exclusionary criteria [6]. A map was created, based on a grid file, for each suitability criterion, and a final composite map was then produced by overlapping all individual grids. The result is a composite map, which highlights the areas that satisfy all suitability criteria (Fig. 2).

Fig. 1. Location map showing the study area.

The GIS software used for the analysis of all vector data is the MapInfo Professional ver. 8.5. Raster data processing and grid creation and analysis were performed using Vertical Mapper ver. 3.1. Topographic, demographic, land use and hydrologic data for the study area is part of the GIS database GR Survey – Digital Data for Greece [7].

The final step in the present work was the estimation of the required area for stabilization pond systems based on the mean minimum monthly temperatures of record, the population of each municipality and the required wastewater effluent discharge criteria, according to the methodology proposed by [8], [9]. The required areas of each municipality were compared to the suitable areas that resulted from the GIS analysis, in order to accept stabilization pond systems as a solution to wastewater treatment problem for each one of the examined municipalities.

For each variable mentioned, a grid file was created for use in the analysis. The whole process was divided in two parts: The first dealt with the creation of a grid for each variable while the second dealt with grid analysis. To increase the speed of the final analysis, for each variable, the created grid showed only the areas of interest, i.e., those that satisfied the criteria for the specified variable. In that way, the rest of the area in the study region could be ignored and was not processed. The scale of capture of the topographic data was 1:250 000 and the Digital Elevation Model produced for the study area had a cell size of 60m x 60m. The same cell size was adopted for all other grids for the other variables, described below.

The slope grid: For stabilization pond systems to have an effective operation, the area should have a relatively mild topography. Steeper areas would not be economically appropriate because they would require excessive excavation. Thus, a 5% slope was considered as the maximum slope value. Areas with higher slope values were excluded from further consideration, as not appropriate for stabilization pond system siting.

The hydrogeological and tectonic setting: In order for stabilization pond systems to have the minimum impact on the groundwater resources, it was necessary to examine the hydrogeological setting of the study area. The process involved digitization of the hydrogeological map 1:200 000 provided by the Greek Geological Survey and the identification of two major categories: those that were or had the potential to become groundwater resources (aquifers) and those that undoubtedly could be considered as aquitards, with no potential as groundwater resources. Areas with little or no data available were assigned to the first category (aquifers), in order to minimize the possibility of aquifer pollution. Areas of the first category were thought as inappropriate for stabilization pond systems and were excluded from the grid analysis.

With regards to seismic activity, the study area shows only minor seismic events and most of it is classified in Zone 1 of the new seismic hazard map of Greece [10]. According to this classification, Zone 1 has a value of active soil acceleration of 0.16g (where g is acceleration of gravity). This is the minimum value for Greece. Only Samothraki island in the study area belongs to Zone 2, i.e., active soil acceleration of 0.24g. Despite the low seismic activity of the study area, buffer zones of 500 meters along both sides of major fault were assigned, so as to prevent the location of the proposed facilities to be on or too close to known faults.

The land use grid: Any potential treatment facility must have as little impact as possible on the existing population. With this in mind there were only two land use classifications distinguished as acceptable for stabilization pond treatment systems: non-forested areas, which are agricultural or populated areas and grasslands. The rest of the land use areas were rejected, as they were covered by sparse or dense forest. In order to include the land use variable in the final analysis, a land use grid was created based on a land use map of 1:500000.

The distance to existing cities and villages: In order to avoid any inconvenient effects to the existing populated areas due to the presence of a wastewater treatment system, it is necessary to locate such facilities at least 500 meters away from villages or city limits. The Greek law does not specify minimum distances from city limits for wastewater treatment facilities. This distance was then selected based on the [11] for sanitary landfills. Although the proposed facilities do not imply as serious environmental risks as landfill sites, for reasons related to public opposition the same distance of 500 m from residential areas was adopted. Accordingly, a 500-meter buffer was then created around existing populated areas and it was converted into a grid file used in the final analysis.

The distance to existing roads and railways: Again, the Greek law does not specify minimum distances of wastewater treatment facilities from roads and railways. According to the Greek Law for sanitary landfills [11], any proposed site should be at least 500 meters away from highways and railways. In the case of stabilization ponds, a smaller distance of 300 meters from highways and railways was adopted, mainly for visual impacts. The presence of such facilities does not have any significant impact on traffic, compared to the presence of a sanitary landfill, which may involve particle and dust transfer, and increased truck traffic volumes. The 300-meter limitation was not adopted for small provincial roads, since the treatment facility has to have access. A buffer zone of 300 meters was created then on both sides of all highways present in the study area, as well as for the railway. The road and railway buffers were then converted to a grid file, which was used in the final grid analysis.

The distance to major surface water bodies: Streams, rivers, lakes and the coastline offer the main disposal options for effluents after treatment, if an irrigation alternative is not possible. Thus, for practical reasons, the proposed facilities should not be far away from the effluent disposal area. However, according to the EU directives for landfill sites, a 500 m buffer zone should be maintained around significant surface water bodies. In the study area such a zone was created around lakes, marshes and rivers of perennial flow, not only for protecting surface waters from a possible leakage of untreated wastewaters, but also for protecting the facilities from flooding. It should be noted here that although a 500-meter distance away from the major riverbeds seems to be quite reasonable, it does not exclude the flooding risk of an area. For this reason, detailed flooding risk studies should be carried out for each new proposed facility at the finer scale.

The environmental protection status: Thrace is an area of great biodiversity. It contains wetlands of international interest, such as the Delta of River Evros, the Delta of River Nestos, and the lakes Vistonida and Ismarida. Several areas belong to the European network Natura 2000 [12] and should be protected according to the EU conservation policy or national and international treaties, such as the Ramsar Convention of 1971. There also exist national parks and protected areas, such as the National Park of Eastern Macedonia and Thrace. As an example, this park is breeding and nesting place for rare aquatic birds and rest place for migrating species. There are found in the area 307 species of birds, among which 34 are endangered and strictly protected, many species of mammals (20 species), reptiles (31 species), insects (110 species), and rare vegetation. The later includes riparian forests, such as the riparian forest along the Delta of River Nestos, and freshwater and coastal wetland species, among which eight are rare and/or endangered [13], [14].

Natural wastewater treatment systems such as stabilization ponds may well exist within protected area boundaries, but only in zones of low protection; they cannot be installed in areas having the maximum protection (e.g., Zone A of the National Park of Eastern Macedonia and Thrace). Thus, these regions were excluded. A grid file was created then for the areas mentioned above, for use in the final analysis.

2.1 Analyzing the Variables

Having all the different variables examined in the study in grid format, the final analysis was carried out using the GridQuery function provided by the raster GIS program Vertical Mapper ver. 3. The program analyses each one of the 60 x 60 meter grid cells for each one of the variables mentioned in the previous sections. If all the predefined suitability criteria are met then the particular grid cell is accepted for use as stabilization pond treatment system. The result is a grid file showing the areas that meet all the previously specified criteria and are suitable for natural treatment systems, such as stabilization ponds (Fig. 2).

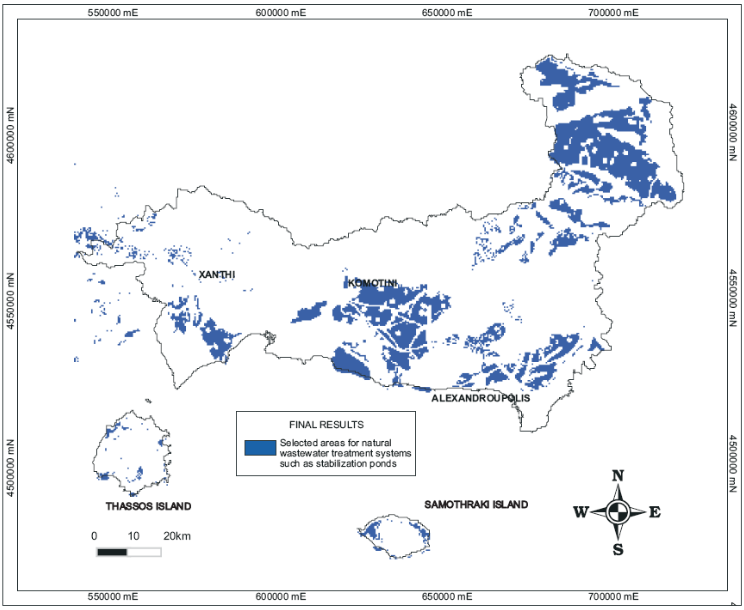


Fig. 2. Grid presenting selected areas for stabilization ponds.

2.2 Sizing and Performance Criteria of Natural Wastewater Treatment Systems

In this final part of the study the calculations are focused on each one of the 36 municipalities of the area. In order to calculate the required area for stabilization ponds in each municipality, the graphical solutions provided by Economopoulou and Tsihrintzis [8], [9] were adopted. As mentioned, these solutions are based on theory by Mara and Pearson [15], [16], and Mara et al. [17]. The graphical solutions calculate the unit area requirements (m^2/capita) of stabilization ponds designed based on BOD_5 and total coliform removal criteria, as a function of wastewater temperature, population served and influent quantity and quality, i.e., weak, typical or strong municipal wastewaters [8], [9]. In this study a typical municipal wastewater was used, i.e., the influent had the following characteristics: unit flow 150 L·capita^{-1}·d^{-1}, BOD 330 mg L^{-1}, total coliform number 10^8/100 mL. The stabilization pond systems are designed to meet two effluent quality characteristics (performance criteria):

- Performance criterion I corresponds to a BOD effluent concentration of 30 mg L^{-1} (e.g., for effluent discharge into small rivers) [18];

- Performance criterion II corresponds to a total coliform effluent concentration of 1000/100mL (e.g., for effluents used for irrigation or discharged near the coastline in bathing waters, according to the Greek Official Gazette [18]). As far as this performance criterion is concerned, other standard values of coliform effluent concentration are also proposed in the relevant literature; for example, Tsagarakis et al. [19] suggest 1000 fecal coliforms/100mL for irrigation of crops and 10000 fecal coliforms/100mL for irrigation of wooden areas or discharge into surface water bodies where public contact is not allowed. Thus, the value of 1000/100mL total coliforms in the effluent, which also agrees with current local government decisions for the Region of Thrace [18], seems reasonable. Finally, this value is also suggested by WHO [20].

For these two performance criteria and influent characteristics, the graph of Figure 3 gives unit areas (m^2/capita) for a stabilization pond system comprising an anaerobic, a facultative and an optimum number of maturation ponds.

It should be mentioned herein that all lakes in the area are protected by the Natura 2000 network and are areas of the highest environmental protection (e.g., in the National Park of Eastern Macedonia and Thrace) classified as Zone A [[21], and consequently cannot directly receive effluent discharges.

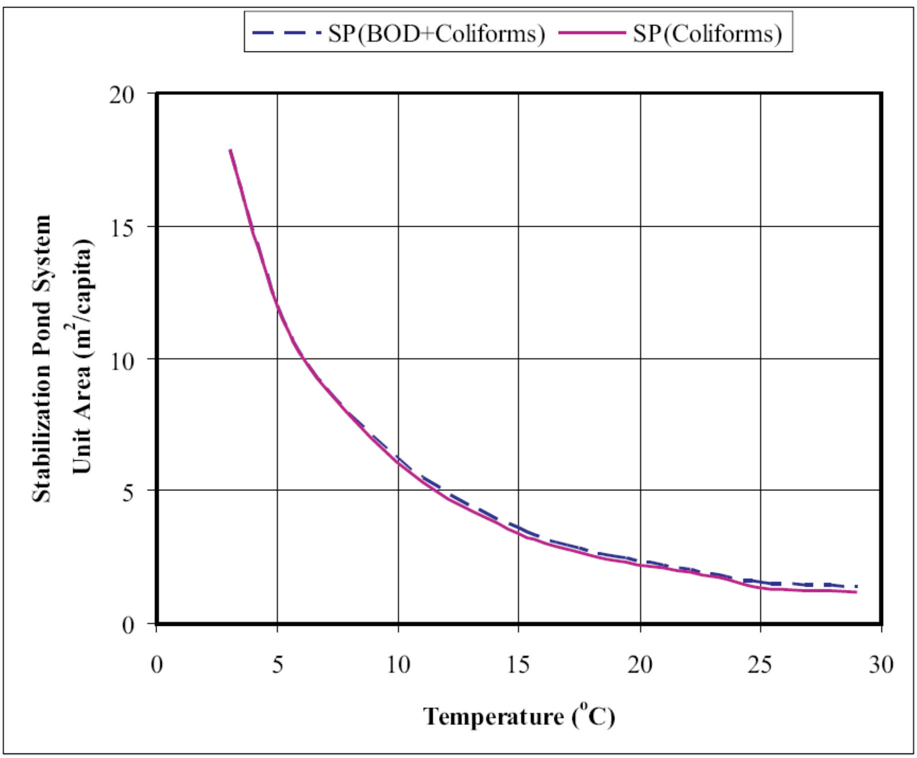


Fig. 3. Stabilization pond sizing graph.

The sizing process in the present work involved the calculation of the maximum area required in each municipality in order to accommodate all weather conditions. According to Mara and Pearson [15], [16], Mara et al. [17] and Economopoulou and Tsihrintzis [8], the design of stabilization ponds is sensitive to pond temperatures. In the winter, the mean daily pond temperature is warmer by 2 to 3ºC than the mean daily air temperature. Thus, to provide a margin of safety, the design pond temperature should be taken equal to the mean minimum monthly air temperature of record for the coldest winter month for the particular area. Eleven weather stations in the study area provide air temperature records. The time series for the period 1966–2001 was used and the mean minimum monthly air temperature for this period for each station was calculated. In order to compute the mean minimum monthly air temperature for each grid cell in the entire study area, the following temperature–altitude relationship was prepared through regression analysis from the 11 stations:

$$x = (470\text{-}y)/88.5;\ (R^2 = 0.86) \tag{1}$$

where: y = elevation above mean sea level (m); x = mean minimum monthly air temperature (^oC); R^2 = square of correlation coefficient. Using Eq. (1), a grid of the mean minimum temperatures of the study area was created based on the digital elevation model.

The process concluded by computing, from Figure 3, the maximum unit area requirements (m^2/capita) for stabilization pond treatment systems for each municipality, in order to meet the corresponding performance criteria for the minimum temperatures.

Table 1 presents the results of the GIS analysis (columns 1 and 2), showing the fraction of the municipal area (both in surface units and % of the total municipal area) suitable for the specified purpose according to the predefined criteria and the total area required for stabilization pond treatment systems, according to the population of the municipality (column 3). The comparison of the areas in Table 1 (columns 1 and 3) shows, based on the GIS analysis, whether the required area for each municipality is available, in order to accept natural treatment systems, such as stabilization ponds, as a solution to wastewater treatment problem for each municipality. The final results are presented in Fig. 4.

Table 1. Available and required surface area for stabilization pond systems for each municipality.

Municipality name	Available area for stabilization pond system installation based on the GIS analysis (km^2)	Percent of the total municipality area (%)	Total area required (ha) for stabilization pond systems
KYPRINOU	97.3	72.3	4.3
VISSIS	58.2	34.2	10.6
DIDIMOTICHOU	166.0	47.8	23.2
METAXADON	89.0	47.2	7.2

Table 1. (*continued*)

Municipality name	Available area for stabilization pond system installation based on the GIS analysis (km^2)	Percent of the total municipality area (%)	Total area required (ha) for stabilization pond systems
ORESTIADOS	113.3	43.1	23.6
ORFEA	107.0	16.6	11.6
SOUFLIOU	20.3	4.4	12.8
TRIGONOU	130.8	32.9	11.7
ARRIANON	46.4	19.4	9.8
KECHROU	7.8	N/A	N/A
ORGANIS	0.0	0.0	N/A
SAPON	143.7	47.9	11.5
KOMOTINIS*	78.6	22.3	77.2
MARONIAS	113.4	39.2	9.7
NEOU SIDIROCHORIOU	10.8	12.5	3.7
FILLIRAS	42.4	17.2	13.1
AMAXADON	0.9	2.5	2.5
AVDIRON	55.0	34.2	4.5
VISTONIDOS	12.8	8.0	10.2
THERMON	0.0	0.0	N/A
KOTILIS	0.0	0.0	N/A
SATRON	0.3	0.2	1.9
SELEROU	0.3	1.1	5.5
EGIROU	16.1	8.5	5.0
IASMOU	2.2	1.0	9.2
SOSTOU	0.2	0.1	11.9
MIKIS	4.4	N/A	N/A
XANTHIS*	9.2	6.0	75.2
STAVROUPOLIS	27.8	N/A	N/A
FERON	87.3	21.4	11.8
ALEXANDROUPOLIS*	108.0	16.8	70.3
SAMOTHRAKIS	31.2	17.3	5.5
TRAIANOUPOLIS	26.5	16.2	4.2
TICHEROU	9.2	4.2	5.6
TOPIROU	27.9	9.0	15.6
THASOU	28.0	7.3	22.7

N/A: The method is not suitable for the specified municipality due to the low minimum temperatures.

* Cities with existing conventional wastewater treatment systems. Stabilization pond systems can be used for upgrading existing treatment plants.

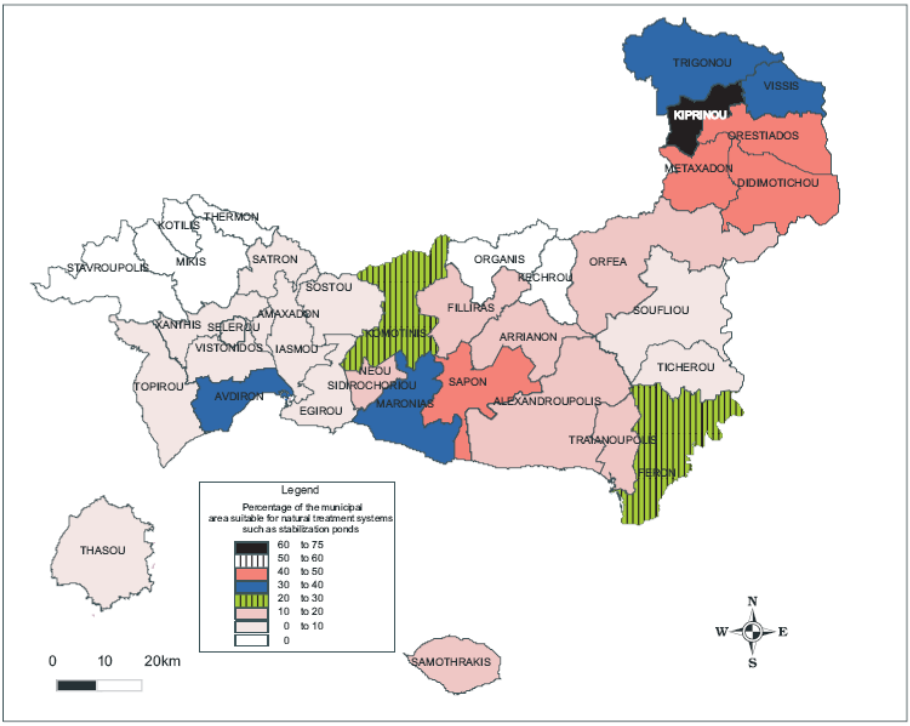


Fig. 4. Municipalities where natural wastewater treatment systems, such as stabilization ponds, are a viable solution.

2.3 Discussion

The method applied derives from the coupling of the GIS technology and a sizing methodology for SP treatment systems using design methodology and effluent discharge criteria. The GIS technology was used in order to create and analyze several grids of different themes and then highlight the areas of interest, i.e., those that satisfy all the relevant criteria. The sizing and performance method calculates the area requirements based on BOD_5 and total coliform effluent discharge criteria as a function of temperature and population. The combination of the above two methodologies, at the municipal level, shows that it is quite simple to check whether natural treatment systems, such as SPs, can be an alternative wastewater treatment option for each one of the examined municipalities and to calculate the maximum area required for such facilities.

In conclusion, it is believed that the present study offers a quite simple and fast way to examine large areas and to highlight possible locations for installation of natural treatment systems, such as SPs, using only exclusionary criteria, as the application of non-exclusionary criteria would require the adoption of finer scale data, more complex computations, more time, and consequently could not be appropriate for covering a large area at a master plan level. However, the final decision, at a municipal level, should be based on more detailed examination on each selected site, involving ranking all possible sites using non-exclusionary criteria, and taking into account the opinion of the local communities, which play perhaps the most important role in siting such facilities.

3 Weighted Linear Combination — Combining GIS, Multicriteria Evaluation Techniques and Fuzzy Logic in Siting MSW Landfills

Siting a MSW landfill is a challenging task, as most of the times various controversial parameters should be considered. GIS can offer significant help in handling this problem. The Analytical Hierarchy Process (AHP) [22], [13] has been incorporated in the GIS technology producing a flexible way of combining various criteria in the siting process [4], [6]. Several studies have been focused on this subject, including evaluation of many factors and aggregation of these factors in many different ways [23], [24], [25], [6], [26], [27], [28].

This section presents a method of determining site suitability for construction of MSW landfills, using the AHP, combined with Fuzzy Function Standardization and Ordered Weighted Averaging (OWA) techniques [3], in a GIS environment. An attempt has been made to apply the OWA technique in a real case, and the results of this relatively new and experimental technique are presented in the following paragraphs.

The procedure analyzes criteria of two types, i.e., constraints (exclusionary criteria) and factors (non-exclusionary criteria). Constraints limit the analysis to particular geographic areas, i.e., they differentiate areas that can be considered suitable for landfill siting by assigning an index value of 1 to those that in no case can be considered suitable and receive an index value of 0. In contrast, factors are criteria that define some degree of suitability for all geographic regions. They define areas or alternatives in terms of a continuous measure of suitability. Individual factors may either receive high scores, thus enhancing, or receive low scores, thus detracting from the overall suitability of an alternative [3]. The aggregation procedure plays a crucial role on the degree each factor influences the final suitability determination. Factors can be standardized in a number of ways depending upon the individual criteria and the form of aggregation used.

In the study eighteen criteria are examined, based on the relevant international literature and were adapted to the Greek and European Union law. They were divided into three main groups, according to how they are considered to influence the landfill site suitability: (1) Constraints; (2) Factors relevant to environmental concerns; (3) Factors relevant to socioeconomic and design parameters. Constraints do not participate in the weight assignment process, whereas all factors in groups 2 and 3, are assigned factor weights, according to their relative importance, based on the AHP [22].

Besides assigning weights to factors, a second set of weights, order weights, has been applied for the different rank-order positions of factors at every location (pixel). The order weights first modify the degree to which factor weights influence the aggregation procedure [3], thus they govern the overall level of risk and trade off in the suitability determination.

Nine constraining criteria are evaluated: (1) residential areas (RA); (2) land uses (LU); (3) highways and railways (HR); (4) environmentally protected areas (EPA); (5) important aquifers (IA); (6) surface water bodies (SWB); (7) springs and wells (SW); (8) exceptional geological conditions (GEO); and (9) distance from country borders and the coastline (DCB). Three factors relevant to the environment are analyzed: (1) hydrogeology (HDG); (2) hydrology (HDR); (3) distance from water bodies (DWB). Six factors relevant to socioeconomic and design parameters are examined:

Fig. 5. Location map showing the study area.

(1) proximity to residential areas (PRA); (2) site access (SA); (3) type of land use (TLU); (4) proximity to waste production centers (PWPC); (5) site orientation (SO); and (6) slope of the land surface (SLS).

The methodology was applied to the prefecture of Evros, located in the northeast continental part of Greece (Fig. 5). The region has an area of 4,000 km^2. It is a rural part of the country, with a population of 151,000 people (according to census data provided by the Greek National Statistical Service for the year 2001) distributed in villages and small towns all over the area. The projected population for 2020, using an average annual population increase of 0.5%, is 161,000. Agriculture and stock farming are the main activities of the population. Administratively, the study area is divided into 13 municipalities, each currently having at least one non-sanitary landfill, while still there is no legal MSW landfill in Evros prefecture.

The study area was discretized using a grid cell size of 30 m x 30 m. Initially all factors in groups 2 and 3 were quantified in a 0 to 255 byte scale, using fuzzy membership functions, and were introduced into the process as map layers. This particular scaling is used because it provides the maximum differentiation possible while analyzing data in byte type, requiring thus half the computer disk space needed for the normal two-byte integer files [3]. Zero is assigned to the least suitable areas for landfill siting and 255 to the most suitable ones, transforming the different measurement units of the factor images into comparable suitability values. In this process, sigmoidal ("s-shaped") fuzzy membership functions, specified for each factor, are used, i.e., monotonically increasing, monotonically decreasing or symmetrical. The sigmoidal membership function is perhaps the most commonly used function in Fuzzy Set theory [3], offering a gradual variation from nonmembership, i.e., 0, to complete membership, i.e., 1. The sigmoidal membership function, can be specified by four parameters (a: membership rises above 0; b: membership becomes 1; c: membership falls below 1; d: membership becomes 0). It is expressed as:

$$\mu(x) = \cos^2 a \tag{2}$$

where, in the case of a monotonically decreasing function:

$$a = \frac{x-c}{d-c} \cdot \frac{\pi}{2} \tag{2a}$$

When x < c, $\mu(x) = 1$

In the case of a monotonically increasing function:

$$a = \frac{1-(x-\mathrm{a})}{b-\mathrm{a}} \cdot \frac{\pi}{2} \tag{2b}$$

When x > b, $\mu(x) = 1$.

Constraining criteria of group 1 remain as Boolean images, dividing the study area in two land categories, i.e., suitable (suitability index 1) and unsuitable (suitability index 0).

The mathematical formulation for area selection, using exclusionary criteria only, is:

$$\mathrm{SI} = \prod_{j=1}^{K} b_j \tag{3}$$

where: SI = overall suitability index value (0 or 1); b_j = suitability index value for each constraining criterion (0 or 1); K = number of constraining criteria.

In each factor group, weights were assigned according to how important each factor is considered. The mathematical formulation for the assignment of the overall suitability index, applying both exclusionary (constraints) and non-exclusionary (factors) criteria, is:

$$\mathrm{SI} = \sum_{i=1}^{N} w_i x_i \cdot \prod_{j=1}^{K} b_j \tag{4}$$

where: SI = overall suitability index value; w_i = weight of factor i; x_i = criterion score of factor i; b_j = criterion score of constraint j; N = number of factors; K = number of constraining criteria.

To make the process of assigning factor weights more objective, a pairwise comparison is applied, in which only two criteria are considered at a time; thus, it becomes more likely to produce a robust set of criteria weights [3]. The implemented technique of factor pairwise comparison was developed by Saaty [22] in the context of a decision making process known as the AHP [13]. In the procedure of multicriteria evaluation in the present study, using the weighted linear combination described above, it is necessary that the assigned factor weights sum to one. In Saaty's [22] technique, factor weights can be derived by taking the principal eigenvector of the square reciprocal matrix of pairwise comparisons between the criteria, as shown in the example presented in Table 2, where the weight assignment process for a six factor problem is presented. Only the lower-left triangular half is actually filled in, since the upper-right triangular half is equal to the reciprocal of the corresponding entries in the lower-left. The cells in the matrix contain the evaluation for each possible pair of comparisons. To rate each pairwise comparison and to fill in the matrix cells, one moves from column to column from left to right. In each cell the relative importance of the row variable to its corresponding column variable is considered. Variables are rated according to the 9-point scale shown in Table 2.

Table 2. Pairwise comparison matrix for a six-factor problem

	PRA	SA	TLU	PWPC	SO	SLS	Calculated factor weights
PRA	1	1	1/3	3	3	1	0.1623
SA	1	1	1/3	1	5	1	0.1506
TLU	3	3	1	5	3	3	0.3784
PWPC	1/3	1	1/5	1	3	1/3	0.0920
SO	1/3	1/5	1/3	1/3	1	1/3	0.0545
SLS	1	1	1/3	3	3	1	0.1623

9: relative to the column variable, the row variable is extremely more important
7: relative to the column variable, the row variable is very strongly more important
5: relative to the column variable, the row variable is strongly more important
3: relative to the column variable, the row variable is moderately more important
1: relative to the column variable, the row variable is equally important
1/3: relative to the column variable, the row variable is moderately less important
1/5: relative to the column variable, the row variable is strongly less important
1/7: relative to the column variable, the row variable is very strongly less important
1/9: relative to the column variable, the row variable is extremely less important

Consistency ratio = 0.07

Order Weights:	0.175	0.170	0.165	0.165	0.165	0.160
Rank:	1st	2nd	3rd	4th	5th	6th

The comparisons concern the relative importance of the two criteria involved in determining suitability for the stated objective [3]. An index of consistency, known as consistency ratio (CR) [22], can be produced, since the complete pairwise comparison matrix contains multiple paths by which the relative importance of criteria can be assessed, thus determining the degree of consistency that has been used in developing the ratings. The consistency ratio indicates the probability that the matrix ratings are randomly generated. Saaty [22] indicates that matrices with CR ratings greater than 0.10 should be re-evaluated. In addition to the overall consistency ratio, it is also possible to analyze the matrix to determine where the inconsistencies arise [3].

Additionally, one more set of weights, order weights, has been applied to each group of factors analyzed, following a procedure known as Ordered Weighted Averaging [3]. Order weights are quite different from factor weights. They do not apply to any specific factor. Instead, they are applied on a pixel-by-pixel basis to factor scores as determined by their rank ordering across factors at each location (pixel). Order weight 1 is assigned to the lowest-ranked factor for that pixel (i.e., the factor with the lowest score), order weight 2 to the next higher-ranked factor for that pixel, and so forth. In that way, a single order weight can be applied to pixels from any of the various factors depending upon their relative rank order. This procedure offers the possibility to adjust the computations, according to how risky or how strict the decision is desired to be. Taking as an example factor group 2, including factors related to the environmental criteria, i.e., hydrogeology (HDG), hydrology (HDR), distance from water bodies (DWB), the assignment of order weights could have the results shown in Table 3.

Table 3. Example of order weight assignment

Moderate Level of Risk – Moderate Strict Decision – Full trade off			
Factors	HDG	HGR	DWB
Order Weights	0.33	0.33	0.33
Rank	1st	2nd	3rd
Low Level of Risk – Very Strict Decision – No trade off			
Factors	HDG	HGR	DWB
Order Weights	1	0	0
Rank	1st	2nd	3rd
High Level of Risk – No Strict Decision – No trade off			
Factors	HDG	HGR	DWB
Order Weights	0	0	1
Rank	1st	2nd	3rd

In the first case, weight is distributed evenly among all factors regardless of their rank order position. The result is moderate in terms of risk, allowing the full trade off between factors, and is placed in the middle between the Boolean AND and OR operation. It is identical to the result produced by assignment of factor weights only and no order weights. In the second case, the factor with the lowest score, i.e., the less suitable, receives all the weight, regardless of the factor weights assigned. The result incorporates the lowest level of risk and no trade off, as all the pixels in the study area were assigned the value of the lowest suitability factor, thus limiting to a high extent the suitable areas. The result will closely resemble the AND (minimum) operation used in the constraining criteria. In addition, because no rank order position other than the minimum is given any weight, there can be no trade off between factors and the minimum factor alone determines the final outcome. On the contrary, the third case, assigns all weight to the factor with the highest suitability score in each pixel, resulting in pixel values equal to the highest suitability factor. The result is very similar to the Boolean OR (maximum) aggregating operation.

Skewing of Order Weights towards either factor with the lowest or highest score results in all possible cases between risk and trade off, or between the AND and OR operation. Altering order weights in terms of their skew and dispersion results in an almost infinite range of possible solutions [3] to the MSW landfill siting problem.

When analyzing a problem, such as sanitary landfill siting, it is clear that the factors involved in the problem do not have the same level of trade off. Factors relevant to the costs of landfill construction and operation are targeted towards minimizing economic costs and can fully trade off with each other. In this way, a high construction cost due to steep slopes of the land surface, can easily trade off with low operational costs due to minimized haul distances, in case a location closer to the waste production centers is selected. On the contrary, when environmental factors are analyzed, it is not that clear to which extent they can trade off, as, for example, locating a landfill as far as possible from surface water bodies does not compensate for a polluted aquifer. This was the main reason to decide to divide all factors involved to two distinct factor groups and treat each group separately in terms of levels of trade off, according to the methodology proposed by Eastman [3].

In order to apply the above described methodology all factors were introduced as map layers in the GIS program MapInfo Professional ver. 8.5. Raster data were processed with Vertical Mapper ver. 3.1, whereas Fuzzy Factor Standardization and the multicriteria evaluation have been performed using the GIS program Idrisi Kilimanjaro [3].

3.1 Criteria Group Description and Application

Constraints

Constraints limit the alternatives under consideration and are expressed in the form of a Boolean (logical) map. Areas which are excluded from consideration are assigned a value of 0, while those open for further consideration are assigned a value of 1 (Fig. 6).

Residential areas: According to Greek law [11], a sanitary landfill cannot be located within 500 m of residential areas. Therefore, a 500 m buffer zone was applied to existing urban areas and industrial areas. For safety, a 3 km buffer zone was applied around airports [29], [6].

Land uses: In the study area agricultural land uses prevail. In the mountainous areas forests of various types are present, which were excluded from further consideration. In order to detect suitable areas for MSW lanfill siting, forest inventory and census data from the GIS database “Gr Survey – Digital Data for Greece” [7] were combined to Landsat ETM satellite images from June 2000. The Normalized Difference Vegetation Index (NDVI) image was created, which expresses the contrast between red and near-infrared reflectances of vegetation and is indicative of the abundance of pigments such as chlorophyll, or simply leaf area. The NDVI is expressed on a scale between -1 to +1, ranges between -0.2 and 0.1 for snow, inland water bodies, deserts and exposed soils, and increases from about 0.1 to 0.7 for increasing amounts of vegetation. In the case of dense leaf canopies, e.g., the humid tropical forests, the NDVI saturates [30]. Finally, negative values represent non-vegetated surfaces [3].

Combining these sources of information offered the possibility to distinguish two discrete types of land uses: unsuitable for landfill siting, i.e., those comprising dense and sparse forests, which received a suitability index of 0, and suitable for further consideration, i.e., those comprising mainly agricultural and uncultivated land which received a value of 1. Agricultural land uses were not excluded from consideration, although they were assigned a low suitability score while analyzing the factor type of land use. This way the probability of interfering landfill operation with agricultural activities was minimized, in case a better alternative existed, i.e., siting sanitary landfills in non-cultivated areas.

Highways and railways: A 500 m buffer zone was maintained on each side of highways and railways. Roads other than highways were not treated as constraining features, since the selected landfill sites should have access.

Environmentally protected areas: The study area contains wetlands of international interest, such as the Delta of River Evros and Dadia National forest. Several areas belong to the European network Natura 2000 [12] and should be protected according to Greek law, EU conservation policies or international treaties, such as the Ramsar Convention of 1971. According to the Greek law [11], landfill siting within areas of monumental or recreational value is forbidden. In addition, a buffer zone of 500 m was maintained around protected areas, designating them as unsuitable for landfill siting.

Important aquifers: A landfill site should be located in an area with low groundwater pollution risk. In order to assess groundwater vulnerability to pollution, the methodology presented by Gemitzi et al. [31] has been applied. According to this study, five discrete groundwater vulnerability classes are designated, based on intrinsic aquifer parameters and on external to the aquifer system parameters. In the present study only intrinsic aquifer parameters are examined. External parameters, such as land uses or surface runoff accumulation, are also incorporated as individual criteria within the multicriteria process presented herein. The intrinsic aquifer parameters are: type of aquifer (i.e., confined, unconfined, fractured, karstic), depth to the water table, and hydraulic conductivity. Based on these parameters a groundwater vulnerability to pollution map is created (see section 1.2.2), with values ranging from 0 (least vulnerable areas) to 255 (most vulnerable areas), and it was categorized into five discrete classes. Areas belonging to the fourth (high groundwater vulnerability to pollution) and fifth groundwater vulnerability class (very high groundwater vulnerability to pollution), i.e., vulnerability index greater than 150 were considered as unsuitable for landfill siting.

Surface water bodies: According to the EU directives, a 500 m buffer zone should be maintained around significant surface water bodies. In the study area such a zone has been created around lakes, marshes and rivers of perennial flow.

Springs and wells: According to the EU directives, landfills must not be close to any source of water, due to pollution risk. Kontos et al. [6] suggests a 500 m buffer zone, which corresponds to a 50 to 60-day groundwater travel time required for pathogen die-off, considering an average groundwater velocity of 10 m/d. The same

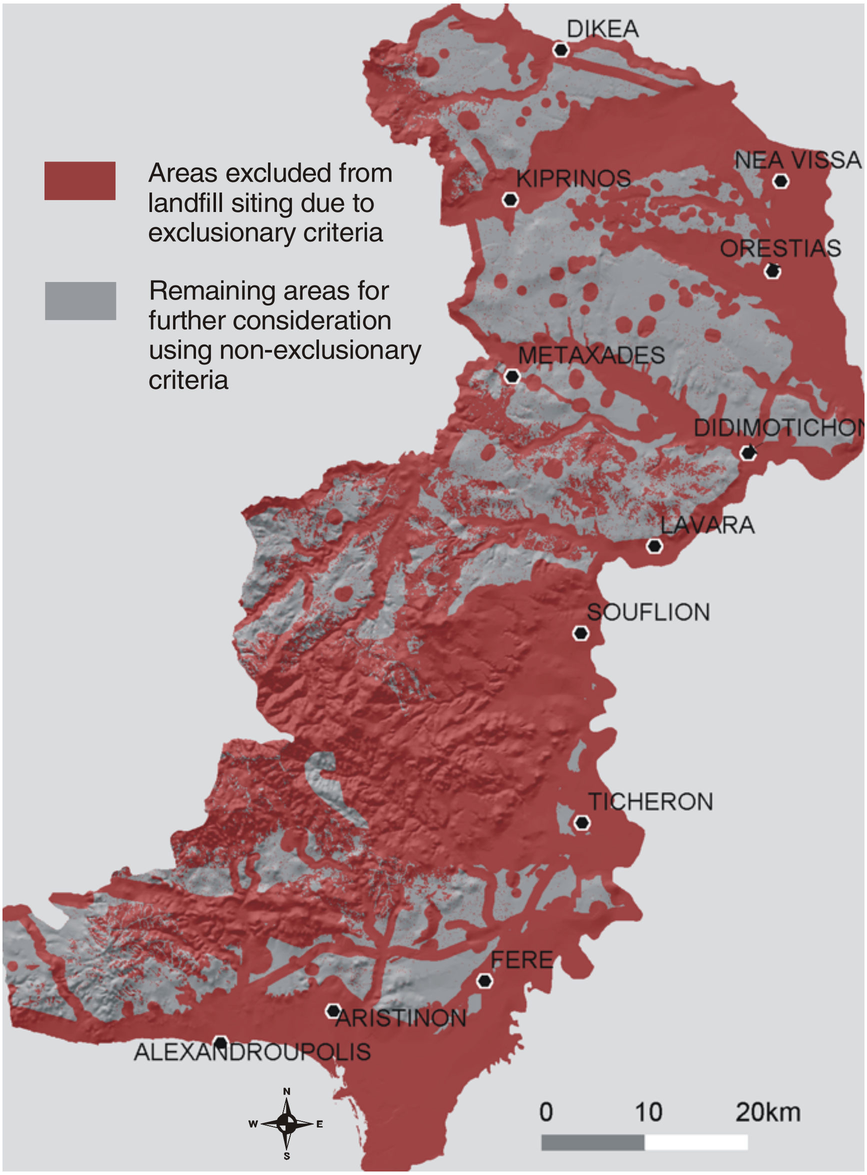


Fig. 6. Areas excluded from landfill siting due to exclusionary criteria

buffer zone is maintained in the present study around sources of drinking or irrigation water, i.e., springs and wells.

Exceptional geological conditions: Presence of active faults, geothermal fields and salt water intrusion zones are prohibitive geological conditions for siting a sanitary landfill. Despite the low seismic activity of the study area, there are active faults in the southern part of the study area, as shown by the location of recent seismic epicenters, which caused earthquakes of small to medium magnitude. For this reason, buffer zones of 1000 meters along both sides of active faults were assigned, so as to prevent the siting of the proposed facility to be on or too close to known active faults.

Geothermal fields are also present in the study area. Known geothermal fields were delineated, whereas a zone of influence of 500 m around each known geothermal borehole or spring has been created, so as to exclude these areas from the siting process. In the same way, saltwater intrusion areas near the coast were excluded from further consideration, as they already present environmental quality problems.

Distance from country borders and the coastline: The study area borders to the east with Turkey and to the north and northwest with Bulgaria. A buffer zone of 1000 m was maintained from the Greek borders. A protective buffer zone of 3000 m from the coastline was created, to protect not only the coastline but also all the tourist activities taking place in the coastal area.

Factors relevant to environmental concerns

Three factors, i.e., non-exclusionary criteria, related to environmental issues were analyzed: (1) hydrogeology (HDG); (2) hydrology (HDR); and (3) distance from water bodies (DWB).

Hydrogeology: A landfill site should not be located in areas with high groundwater pollution risk. Although the Greek law [11] does not provide specific guidelines, the groundwater pollution risk has been assessed using the methodology for groundwater vulnerability to pollution assessment discussed in the constraining criterion of important aquifers [31], categorizing the study area into five groundwater vulnerability classes (see section 4). In the examination of exclusionary criteria all pixels with groundwater vulnerability value greater than 150 (very high groundwater vulnerability to pollution) were considered unsuitable for landfill siting. The rest of the areas, i.e., pixels with vulnerability value less or equal to 150, were assigned a suitability score, transforming the values from 0 to 150 scale, to a continuous set of values, ranging from 0 to 255. Thus, a monotonically decreasing sigmoidal fuzzy membership function was applied (Fig. 7a).

Hydrology: Surface runoff is a factor that is related both to environmental concerns, as the pollution potential increases in case of storm surges, as well as to economic costs, as it involves the construction costs for an efficient drainage system. Relevant studies examine the area of the upstream hydrologic basin [6], or the floodplain of a specified return period [4], in order to account for this factor. In the present study, a different approach is adopted; each pixel in the study area receives an amount of rain and the specially developed routine, RUNOFF, incorporated in the raster GIS Idrisi Kilimanjaro [3], calculates the accumulation of rainfall units per pixel based on an elevation image (DEM) [32] (Fig. 7b). The higher the accumulation of runoff in a pixel, the higher the flooding risk and the related costs for drainage network construction. A simple RUNOFF analysis accumulates rainfall on a per pixel basis as if one unit of rainfall was dropped on every location. In our case, mean annual precipitation data from 13 gaging stations scattered all over the study area, for the time period 1966 - 2001, was incorporated in the routine. Besides providing a DEM, the routine has been enhanced with a permeability image, in order to account for soil permeability and to adjust the accumulated rainfall accordingly [3].

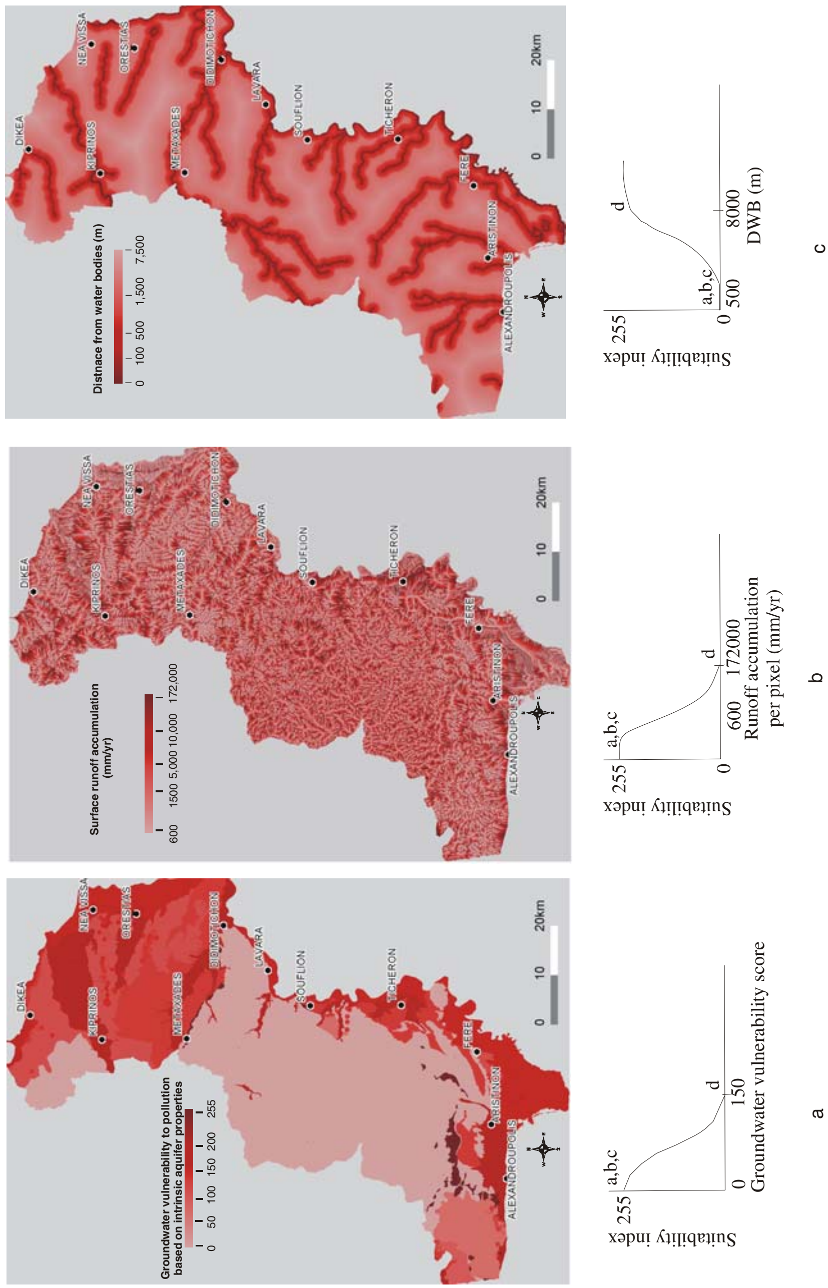


Fig. 7. (a) Groundwater vulnerability to pollution; (b) surface runoff accumulation; (c) distance from water bodies

A monotonically decreasing sigmoidal fuzzy membership function was applied to transform runoff accumulation to a 0 to 255 scale (Fig. 7b). Control point (a) was set to 600 mm/yr where function membership becomes one and control point (d) was set to 172,000 mm/yr where function membership becomes zero. The value of 172,000 mm/yr after multiplication by the pixel size, i.e., 900 m2, corresponds to an average discharge of 425 m3/d. Control points (a) and (d) were selected after imposing a quantiles classification scheme, placing an equal number of pixels of the study area into each class. The first point corresponds to the end of the first class, i.e., 600 mm/yr (approx. 5% of the study area) and the second point corresponds to the starting point of the last class, i.e., 172,000 mm/yr (basin outlets) (upper 5% of the study area) defined as control points (a) and (d), respectively, for the sigmoidal fuzzy function. Pixels with accumulated runoff values lower than 600 mm/yr and higher than 172,000 mm/yr were considered equally suitable to the first and second control points, and received a function membership value of one (highest suitability) and zero (lowest suitability), respectively.

Distance from water bodies: Besides excluding landfill siting in areas within 500 m around significant surface water bodies, i.e., lakes, marshes and rivers of perennial flow, it is preferable to locate such activities as far as possible from surface waters bodies. A location profile has been created, representing the average distance to a series of features from anywhere within a map area [5]. The algorithm generates a grid, where at each cell a value is calculated that represents the average distance to all feature locations, i.e., surface water bodies, surrounding that cell (Fig. 7c). The greater the average distance from surface water bodies, the more suitable the pixel is considered to be for landfill siting.

A monotonically increasing sigmoidal fuzzy membership function was applied in order to transform the calculated distances to a continuous set of values ranging from 0 to 255 (Fig. 7c). Control point (a) was set to 500 m (lowest suitability) where function membership becomes 0, and control point (d) was set to 8000 m (the most distant pixel to water bodies) where function membership becomes 1 (highest suitability).

Factors relevant to socioeconomic and design parameters

Six factors relevant to socioeconomic and design parameters are examined: (1) proximity to residentiall areas (PRA); (2) site access (SA); (3) type of land use (TLU); (4) proximity to waste production centers (PWPC); (5) site orientation (SO); and (6) slope of the land surface (SLS).

Proximity to residential areas: Landfill siting is excluded within 500 m from residential areas. It is desirable, though, to locate a sanitary landfill as far as possible from existing towns, villages or settlements. A location profile was generated, as in the case of the factor DWB (Fig. 8a). This time, however, the weighted average distance of each grid cell of the study area to the five nearest residential areas was calculated. The weighting factor was the population of each residential area according to the 2001 census data, provided by the Greek Statistical Service. The greater the weighted average distance from residential areas, the more suitable the pixel is considered to be for siting a MSW landfill.

A monotonically increasing sigmoidal fuzzy membership function was applied in order to transform the calculated distances to a 0 to 255 scale (Fig. 8a). Control point (a) was set to 500 m (lowest suitability) where function membership becomes 0, and control point (d) was set to 20 km (the most distant pixel to residential areas) where function membership becomes 1 (highest suitability).

Site access: The proposed construction of sanitary landfill would be economically more feasible if it is located close to existing roads. Additional costs for road construction, in areas distant to roads, would make their selection less attractive. A very low suitability value, however, was assigned for pixels within a distance of 100 m from existing roads, in order for landfill vehicles not to interfere with the rest of the traffic. A location profile was created, calculating the distance of each grid cell in the study area to the closest existing road (Fig. 8b).

In this case, a symmetric sigmoidal fuzzy membership function was applied to transform the computed distances to a 0 to 255 scale (Fig. 8b). Control point (a) was located at 99 m (inflection point where function membership increases from 0 to 1), control point (b) was set to 100 m where function membership becomes 1 (highest suitability), control point (c) was set to 250 m (inflection point where function membership starts to decrease) and control point (d) was set to 5 km (the most distant pixel from any road in the study area) where function membership becomes zero.

Type of land use: As mentioned in the constraining criterion LU, all types of forests were excluded from the siting process. Agricultural areas were assigned suitability values according to the NDVI value, which increases from 0.1 to 0.7 for increasing amounts of vegetation, and becomes > 0.7 in the case of dense leaf canopies [30]. A monotonically decreasing sigmoidal fuzzy membership function was applied to transform the NDVI values to a 0 to 255 scale (Fig. 8c). Pixels with lower than 0.1 NDVI values were considered as equally suitable as those with 0.1 NDVI values, whereas pixels with higher than 0.5 NDVI values were assigned a value of zero suitability, even though they were excluded while examining the LU constraining criterion.

Proximity to waste production centers: The proximity to waste production sources represents one aspect of the economic feasibility of a candidate landfill site, decreasing hauling costs by locating landfills near population centers. Taking into account that the present study aims at locating one landfill site for the whole study area, the average distance of each pixel from all the residential centers in the study area was calculated, weighted by the number of inhabitants of each population center (Fig. 8d). This factor is quite different from the PRA, as the former targets at reducing hauling costs, whereas the latter aims at minimizing public dissatisfaction. It should be pointed out herein, that in the case of proximity to residential areas criterion, the weighted average distance from the five nearest population centers has been calculated, as these are expected to be the most dissatisfied ones in case a landfill site is located in their vicinity. On the contrary, in the case of the proximity to waste production centers, the weighted average distance from all residential areas was calculated, as all of them will contribute to waste generation proportionally to their population.

A monotonically decreasing sigmoidal fuzzy membership function was applied to transform the computed distances to a 0 to 255 scale (Fig. 8d). Control point (a) was set to 40 km where function membership becomes 1 (highest suitability), and control point (d) was set to 60 km where function membership becomes 0 (lowest suitability). Pixels with values lower than 40 and greater than 60 received values of 1 and 0, respectively.

Site orientation: The dominant wind orientation in the study area is north to north-east. Aspect measures the direction that each grid cell faces in three-dimensional space, i.e., perpendicular to the strike of a surface, and is recorded in azimuth degrees relative to true north.

Four discrete aspect classes were defined and assigned suitability values ranging from 0 to 255, according to whether they favor odor and particle transport or not. In that way the orientation class from 0° to 90° was assigned a value of 255 (best), the classes from 90° to 180° and from 270° to 360° received a value of 130 (moderate suitability), and the class from 180° to 270° received a value of 0 (unsuitable) (Fig. 8e).

Slope of the land surface: The slope of the land surface is a crucial factor as far as construction costs are involved, as very steep slopes will lead to higher excavation costs. The slope of the land surface was calculated on a pixel basis, as a percentage, ranging from 0% to 50%, using the DEM of the study area (Fig. 8f). A monotonically decreasing sigmoidal fuzzy membership function has been applied to transform the above values to 0 to 255 scale (Fig. 8f). Control point (a) was set to 10% where function membership becomes 1 (highest suitability), and control point (d) was set to 50% where function membership becomes 0 and the associated suitability is diminished. Pixels with value lower than 10% were considered equally suitable for landfill siting as those with slope value of 10%.

Aggregation Procedure

After standardizing all factors to a common 0 to 255 scale using fuzzy membership functions, factor weights were given to all factors in each group. The weights indicate a factor's importance relative to all other factors and control how factors compensate for each other in each factor group.

Factor weights sum to 1 for each factor group that the project has been divided into. Several techniques exist for assigning factor weights. The simplest one could be the division of 1 into the number of factors in each factor group. However, weights produced with this procedure are not often realistic. A more efficient way of producing factor weights is the AHP [22], [13], [33]. In the present study the AHP has been implemented applying the module WEIGHT, incorporated in the raster GIS program Idrisi Kilimanjaro [3]. In this module, each pair of factors in a particular factor group is examined at a time, in terms of their relative importance (Table 2). After all possible combinations of two factors, the module calculates a set of weights that sum to 1 and a consistency ratio. This ratio indicates any inconsistencies that may have been made during the pairwise comparison process, i.e., the probability that factor weights have been assigned quite randomly [3]. A consistency ratio greater than 0.10 indicates that factor weights should be re-evaluated [22].

A Weighted Linear Combination [1] is applied in order to combine information from various factors in each factor group. This aggregation technique multiplies factor scores by their factor weight and then sums the products to yield the suitability score as described by Equation (4).

Nevertheless, when analyzing a siting problem that involves both economic and environmental criteria, it is easy to determine how factors related to economic costs trade off, as they can be converted to monetary value. On the other hand, it is very difficult to decide if and to which extent environmental factors trade off and to convert, for example, the pollution of a water body to monetary cost. Control over risk and trade off is made possible through a second set of weights, order weights, applied to factors on a pixel by pixel basis, for the different rank order positions of factors at every location (pixel). The order weights first modify the degree to which factor weights influence the aggregation procedure, thus they govern the overall level of trade off [34], [35]. After factor weights are applied to the original factors, the results are ranked from low to high suitability for each location. The factor with the lowest suitability score is then given the first order weight, the factor with the next higher suitability score is given the second order weight, and so on. This has the effect of weighting factors based on their rank from minimum (lowest suitability) to maximum (highest suitability) value for each location. The relative skew toward either minimum or maximum of the order weights controls the level of risk in the evaluation [3] (Table 3). Additionally, the degree to which the order weights are evenly distributed across all positions controls the degree to which factor weights influence.

The procedure is repeated once for each group of factors, resulting in two intermediate suitability landfill siting maps. The final suitability map is produced by aggregating, using the same procedure, the intermediate results.

In the following paragraphs the factor and order weights in each one of the two factor groups are presented. In both cases, skewing order weights towards the less suitable factor offers a more conservative solution and stresses pixels in the study area to lower suitability values.

Weighting and Aggregating Factor Group 2: The pairwise comparison matrix and the calculated factor weights for factor group 2 are shown in Table 4, where factor 1 is Hydrogeology, factor 2 is Hydrology, and factor 3 is Distance from water bodies.

Table 4. Pairwise comparison matrix for factor group 2

	HGD	HDR	DWB	Calculated factor weights
HGD	1	1	3	0.4286
HDR	1	1	3	0.4286
DWB	1/3	1/3	1	0.1429
Consistency ratio =	0.0			
Order Weights:	0.7	0.2	0.1	
Rank:	1st	2nd	3rd	

The assigned order weights (Table 4) show a low level of risk and minimized trade off, as environmental factors are not expected to trade off. Skew towards the factor with the lowest suitability score, shows a very strict decision. The intermediate landfill siting suitability map is presented in Figure 9a.

Weighting and Aggregating Factor Group 3: The pairwise comparison matrix and the calculated factor weights for factor group 3 are shown in Table 2.

The assigned order weights (Table 2) are distributed evenly among all factors regardless of their rank order position. Only a minor skew towards factors with the lowest suitability scores makes the decision moderately risk averse, while trade off is maximized, allowing the factor weights to be employed. Figure 9b presents the suitability map for landfill siting based on factor group 3.

Weighting and Aggregating Intermediate Results

Intermediate results were aggregated in order to produce the composite MSW landfill siting suitability map. Constraints always remain as Boolean masks and they are not involved in any weight assignment process. The intermediate suitability maps, which resulted from the two previously examined groups, were used as equally weighted factors in the final aggregation process:

Factor group 1: 0.5; Factor Group 2: 0.5

The following order weights were assigned:

Order Weights:	0.6	0.4
Rank:	1st	2nd

The assigned order weights show a low level of risk. Aggregation is skewed towards the factor with the lowest suitability score.

Sizing procedure

In order to estimate the required landfill area for Evros prefecture, the methodology described by Aivaliotis et al. [36] was adopted. The following assumptions were made: symmetrical combination of landfill and landraise, with base surface area of size A (in 10^3 m^2) and of two shapes, square and orthogonal. The average per capita solid waste production in Greece for the year 2004 has been reported at 1.1 kg/capita/day [36] and the average density in the landfill after compaction is between 500 to 800 kg/m^3 [35]. For the study area, the estimated solid waste quantity M is then 1,309,411 tons for a 20-year operation period, assuming a 0.5% population increase per year and a constant average waste production per capita per year. Thus, the required landfill area ranges from 120,000 to 160,000 m^2, for the two mentioned shapes. For a more conservative approach, the higher value, i.e., 16 ha, has been used in the siting calculations.

3.2 Results

The final suitability results were divided into 5 discrete categories from low to high suitability, i.e., 0 to 50 (very low suitability), 50 to 100 (low suitability), 100 to 150 (moderate suitability), 150 to 200 (high suitability), 200 to 255 (very high suitability), as shown in Figure 9c. Table 5 presents the areas of these classes and the percentages of the total study area. Figure 9d shows the areas of very high suitability.

According to the estimated required area of 160,000 m^2, a post aggregation constraint should be introduced requiring that suitable sites must be at least 160,000 m^2 in size, if only one landfill is to serve the entire study area. This is achieved by selecting a fixed quantity of top-ranked locations, i.e., those that belong to the fifth suitability category (200 to 255), equivalent to the required area. The result is a Boolean map, where exact amounts of land are allocated for MSW landfill (Fig. 10a). Three sites are identified, while their size and average suitability index, i.e., average suitability value of all pixels within the selected site, are presented in Figure 10a.

Moreover, due to the fact that such constructions require permission from military and archeology services, the selected areas should be examined for absence of such features. It is noted that military facility and archaeological data are not available, and therefore cannot be used in the GIS analysis.

A different scenario is also presented in Figures 10b and 10c, where the alternative choice of two sites for landfill construction, i.e., one for the north and one for the southern part of the study area was examined. As each one of these two parts of the prefecture holds almost half of the entire prefecture population, the required area for either of them will be approximately half the predefined required area calculated for the whole prefecture. It should be noted herein that all factors and constraints remained the same, as in the previous aggregation procedure. Only the factor of PWPC changed, as this time only residential areas of either the north or the south part of the study area participated in the calculations. According to this alternative, five areas are designated as suitable, i.e., suitability index $\geq$ 200, in the north Evros region and one in the south Evros region, as shown in Figures 10b and 10c. It should be mentioned that in this work all suitable areas are presented, and not the two most suitable ones. The choice of either scenario and either of the identified sites should also be based on detailed cost benefit studies as well as on public opinion, and is beyond the scope of this study.

Table 5. Landfill siting suitability results

Suitability Class	Suitability index	Area (km^2)	Percentage of the study area
Very low	0 - 50	2720	68%
Low	50 - 100	400	10%
Moderate	100 - 150	680	17%
High	150 - 200	160	4%
Very high	200 - 255	40	1%

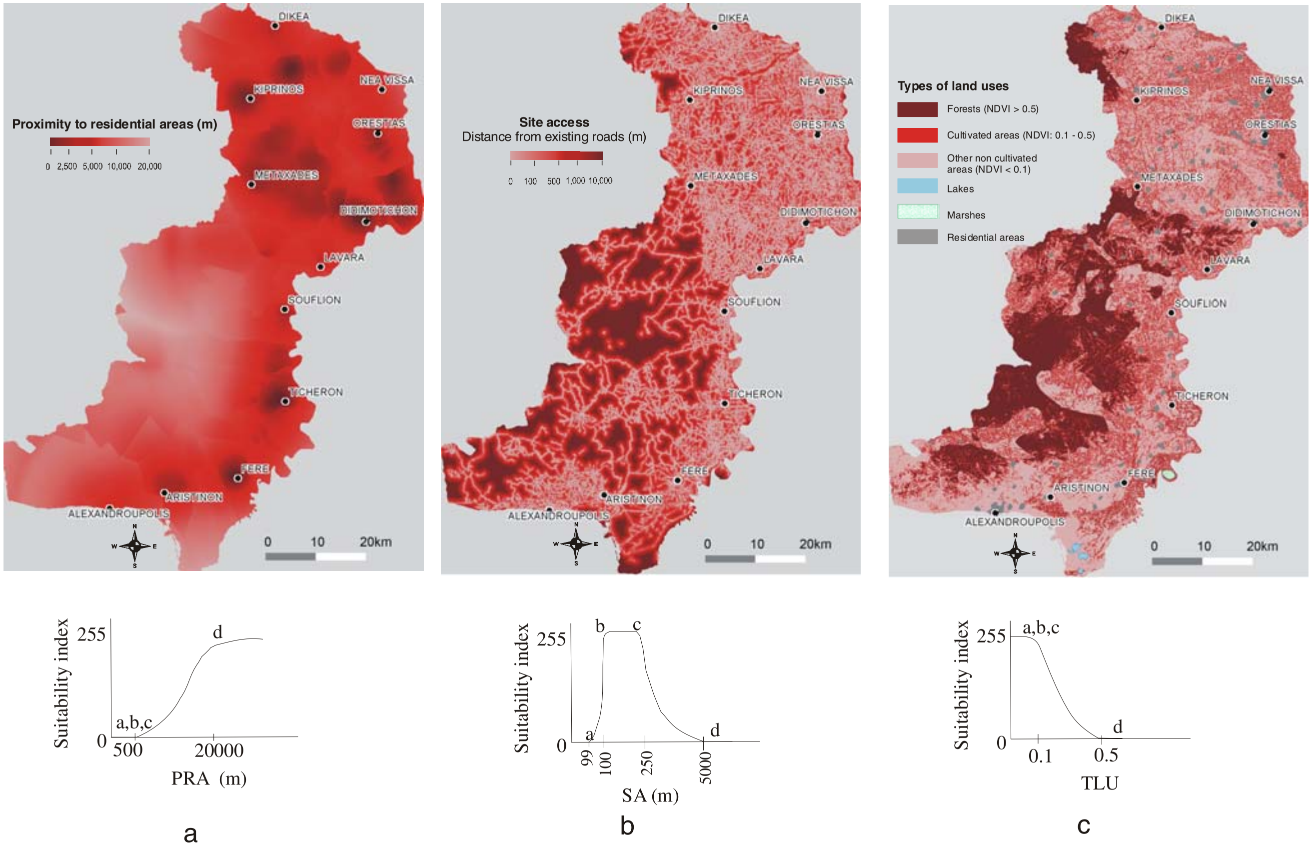


Fig. 8. (a) Proximity to residential areas; (b) site access; (c) types of land uses; (d) proximity to waste production centers; (e) site orientation; (f) slope of the land surface; (e) site orientation; (f) slope of the land surface.

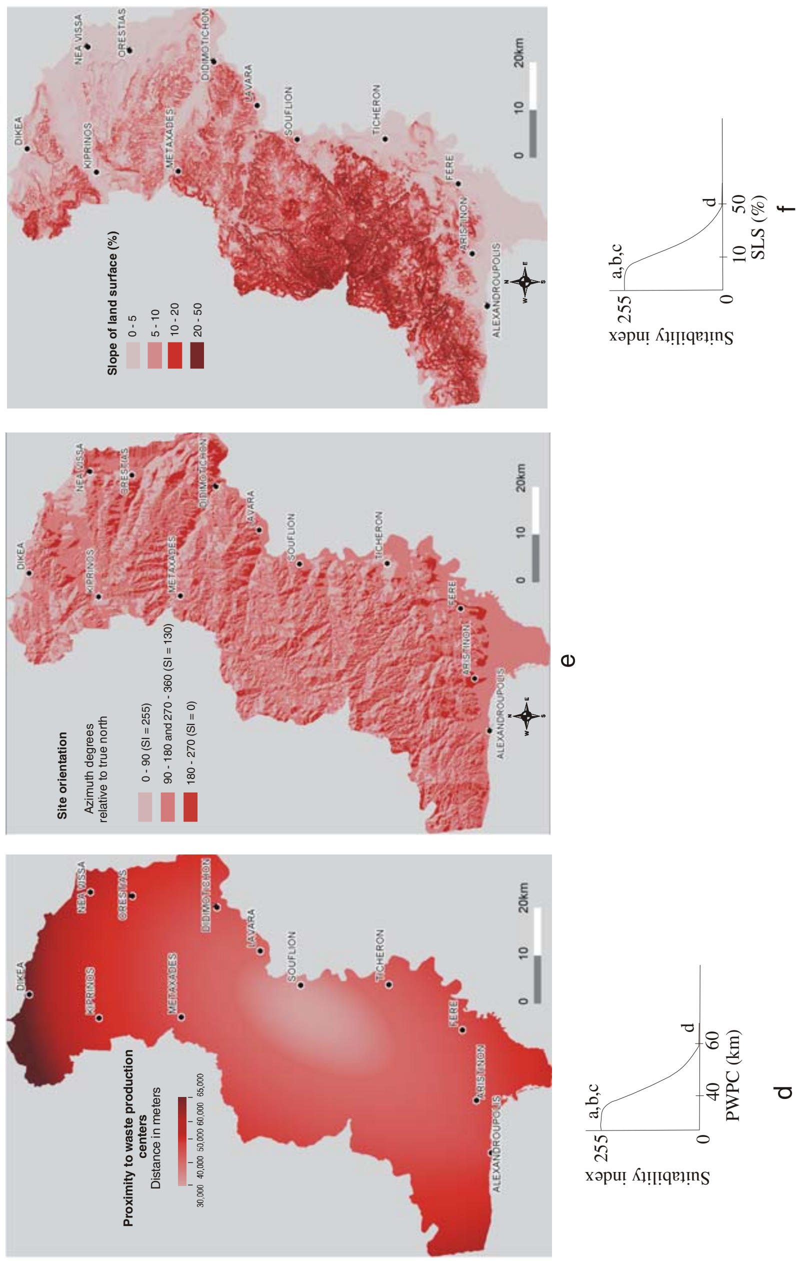


Fig. 8. (*continued*)

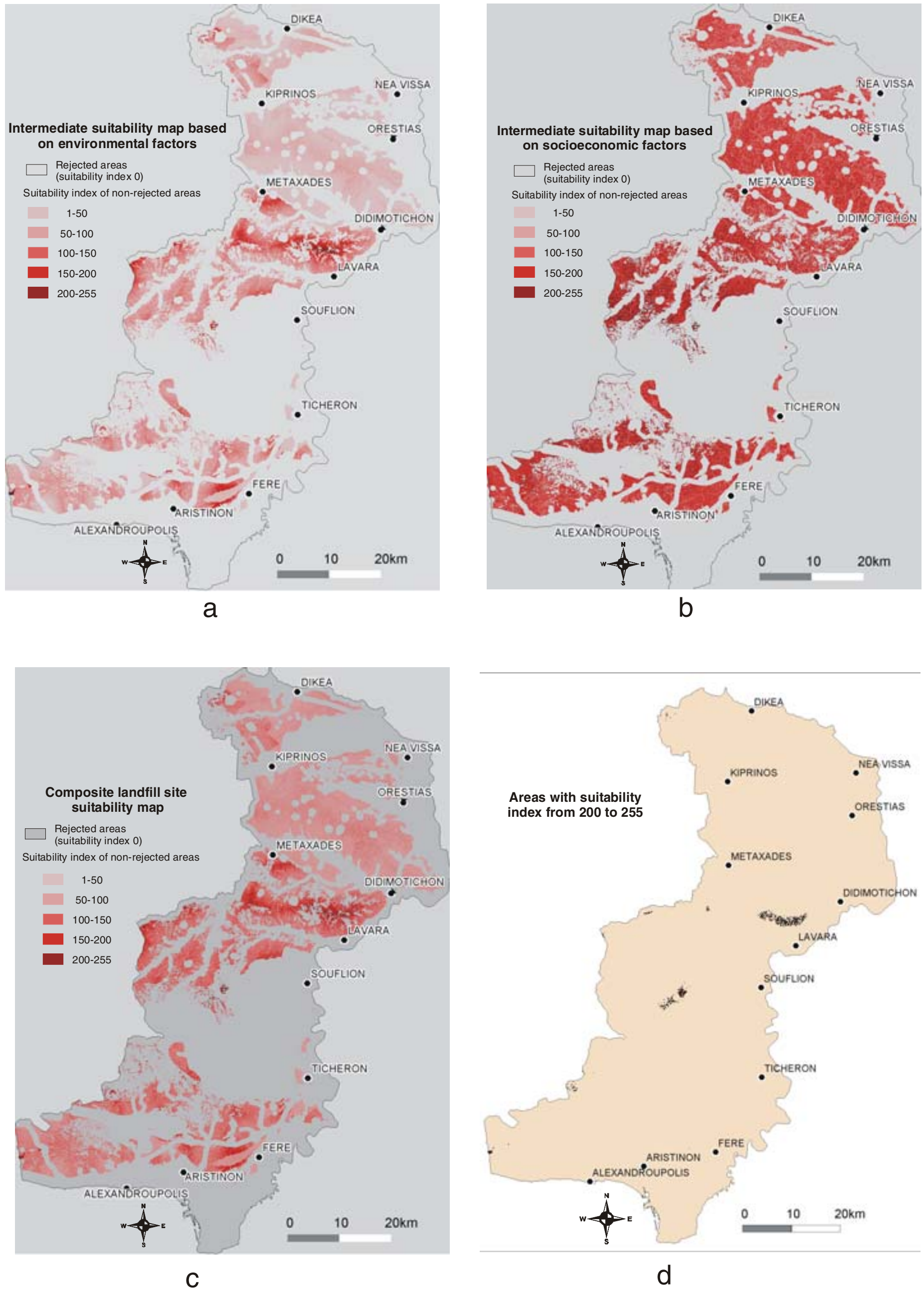


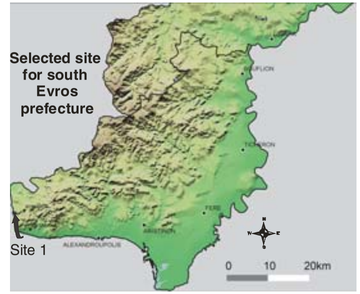
Fig. 9. (a) Intermediate suitability map based on environmental factors; (b) Intermediate suitability map based on socioeconomic factors; (c) Composite landfill site suitability map; (d) areas with very high suitability index.

Selected site	Site area (m^2)	Average Suitability Index
1	224,700	211
2	685,000	213
3	206,000	212

a

Selected site	Site area (m^2)	Average Suitability Index
1	91,060	209
2	171,000	210
3	95,000	207
4	119,000	209
5	111,000	206

b



Selected site	Site area (m^2)	Average Suitability Index
1	250,000	213

c

Fig. 10. (a) Selected suitable sites for Evros prefecture; (b) selected suitable sites for the north part of Evros prefecture; (c) selected suitable sites for the south part of Evros prefecture.

3.3 Discussion

In the present study a methodology for assessing location suitability for MSW landfill was developed that takes into account both exclusionary and non-exclusionary criteria, and coupling GIS with multicriteria evaluation techniques and fuzzy logic. Several approaches have been presented for landfill siting using GIS and multicriteria evaluation [23], [4], [6], [26]. In the present study, however, besides offering all the advantages of the above-mentioned techniques, an important contribution has been achieved through the application of the order weights, on a pixel-by-pixel basis, which offers the full control over the level of risk and trade off desired. Moreover, the distinction of the parameters involved in the siting process into three groups, i.e., constraints which are not included in the weighting process, environmental factors that generally do not trade off and socioeconomic parameters that usually trade off, offers a noticeable flexibility in the aggregation procedure. It is thus quite simple to explore different siting scenarios, as far as a cost-sensitive approach is desired, or a more sensitive environmental alternative is aimed. Furthermore, the choice of the fuzzy membership function control points at which set membership becomes either 0 or 1, was a critical issue in the factor standardization process of the present study. It is important to consider the physical meaning of each variable, before choosing the inflection points of the fuzzy membership function. As far as distances are concerned, inflection points are located at distances specified by the law, or by common sense. In case of factors of broad value range, such as proximity to waste production centers, a quantiles classification scheme was applied, placing an equal number of pixels into each class, and control points were located at the end of the first class and the starting point of the last class (approximately upper and lower 5% of the study area). On the other hand, when analyzing the site access factor, it was merely a choice of common sense the assignment of very low suitability to all areas within a distance of 100 m from any road, for the landfill vehicles not to interfere with other traffic. At that point suitability raises at 100 m, remains very high within distances up to 250 m from roads, where it starts decreasing, while function membership becomes 0 (lowest suitability) at 5 km (the most distant pixel from any road in the study area). This is achieved through the application of a symmetric membership function. Factor weights were assigned to the nine non-exclusionary criteria (factors) involved in the calculation process. It is clear that assignment of factor weights is based on previous knowledge of the factor characteristics and the particularities of the study area, as well as on the experience of the scientists involved in the weight assignment process. It was attempted, however, to develop a weight assigning process as objective as possible, by applying techniques like the AHP. Moreover, applying order weights provided full control over the level of risk and trade off in produced results. In the case of factor group 2, the same factor weight was assigned to hydrology and hydrogeology factors, while most of theorder weight was assigned to the least suitable factor, allowing only a minor level of trade off between factors and producing an environmentally risk averse result. In that way, a very strict decision is made, as far as environmental issues are concerned. In the case of factor group 3, order weights are dispersed to all six factors, leading to a high level of trade off, as socioeconomic factors impose. There is, however, a minor skew toward the least suitability factors, making thus the decision moderately risk averse. In the final aggregation process, factor weights are evenly dispersed to both factors, i.e., environmental and socioeconomic, as they both play an important role while siting a MSW landfill. Order weights are skewed again toward the factor with the least suitability score, making the choice quite risk averse. The prolonged shape of the study area leads to high haul distances, in case a single site is selected for the whole prefecture of Evros. An alternative solution could be the selection of two candidate sites, i.e., one for the north and a different one for the south part of the study area as presented in Fig. 6b, c. The choice of either solution, however, should be based on detailed cost benefit study, which is beyond the scope of the present work. Furthermore, transfer stations can be used to minimize transportation cost.

4 Weighted Linear Combination — Assessment of Groundwater Vulnerability to Pollution: A Combination of GIS, Fuzzy Logic and Decision Making Techniques

Assessment of groundwater vulnerability to pollution aims at highlighting areas at high risk of being polluted. A methodology is presented to estimate the risk of an aquifer to be polluted from concentrated and / or dispersed sources. The methodology applies an overlay and index method, involving several parameters as in the case of section 3. This time, however, no exclusionary criteria are used, as it is not intended to locate potential sites for a particular activity. In addition, the whole study area is examined and assigned groundwater vulnerability to pollution scores. Parameters were categorized into three factor groups: factor group 1 includes parameters relevant to the internal aquifer system properties, thus determining the intrinsic aquifer vulnerability to pollution; factor group 2 comprises parameters relevant to external stresses to the system, such a human activities and rainfall effects; factor group 3 incorporates specific geological settings, such as the presence of geothermal fields or salt intrusion zones into the computation process. Geographic Information Systems have been used for data acquisition and processing coupled with a multicriteria evaluation (MCE) technique enhanced with Fuzzy Factor Standardization. Moreover, besides assigning weights to factors, a second set of weights, order weights, has been applied to factors, on a pixel by pixel basis, thus allowing the control of the level of risk in the vulnerability determination and the enhancement of local site characteristics. Individual analysis of each factor group resulted in 3 intermediate groundwater vulnerability to pollution maps, which were combined in order to produce the final composite groundwater vulnerability map for the study area. The method has been applied in the region of Eastern Macedonia and Thrace (North Greece), an area of approximately 14,000 km^2. The methodology has been tested and calibrated against measured nitrate concentration in wells, in the northwest part of the study area, providing results related to the aggregation and weighting procedure.

Up to now, several vulnerability assessment techniques have been developed. The most common ones are: the DRASTIC system [37], the GOD system [38], the AVI rating system [39], the SINTACS method [40], the ISIS method [41], the Irish perspective [42], the German Method [43] and EPIK [44]. A comparison of the above aquifer vulnerability assessment techniques has been performed by Gogu [45] and Gogu et al. [46], which showed that there is a wide range in the results provided by each method, and that in many cases there is disagreement. The reason for this is that aquifer vulnerability is not a measurable quantity, making the choice among the several methods quite an ambiguous task. In recent studies [47], [48], however, it was attempted to compare the results of vulnerability assessment methods with aquifer water quality data and perform a method sensitivity analysis.

Twelve factors are involved in the computation process, distinguished in three main groups, according to the way they influence groundwater vulnerability to pollution. More specifically, the following twelve factors were introduced into the computation process of aquifer vulnerability: (1) Aquifer type; (2) Depth to water table; (3) Hydraulic conductivity; (4) Surface runoff; (5) Land uses (non-point source pollution); (6) Concentrated polluting activities (point source pollution); (7) Proximity to rivers; (8) Proximity to highways; (9) Proximity to residential areas; (10) Presence of environmentally protected areas; (11) Presence of geothermal fields; and (12) Presence of saltwater intrusion zones. The first group includes factors 1 to 3, relevant to the hydraulic properties of the aquifer, which are not dependent on external stresses, thus characterizing the intrinsic vulnerability of the aquifer. The second group comprises factors 4 to 10, relevant to the socioeconomic and development status of the study area, reflecting the impact of the external anthropogenic forces on the aquifer system, thus describing the specific aquifer vulnerability. In this second group, a factor that describes surface runoff accumulation is also included, as it is considered an external parameter to aquifer system. The third group includes factors 11 and 12, which are relevant to the presence of particular geological conditions, such as areas influenced by saltwater intrusion or those including geothermal fields. In the present study, these three distinct categories were examined separately, producing three types of intermediate aquifer vulnerability to pollution maps.

The study area has been discretized using a grid cell size of 60 m x 60 m. Initially all factors were standardized to a byte-level range of 0 to 255, as in the previous case study of section 3. Zero is assigned to the least vulnerable areas and 255 to the most vulnerable ones, transforming the different measurement units of the factor images, which served as GIS map layers, into comparable values, using fuzzy membership functions. Again, sigmoidal ("s-shaped") fuzzy membership functions, specified for each factor, are used, as expressed in equations (2), (2a) and (2b).

The Analytical Hierarchy Process (AHP) and the Ordered Weighted Average techniques were implemented, as in the case of landfill site selection, for assigning factor weights and order weights, respectively.

For each one of the three factor groups an intermediate aquifer vulnerability map was created; these maps were then combined in two ways in order to produce a composite aquifer vulnerability map. In order to calibrate and check the validity of the presented methodology, the results were compared to measured nitrate concentration in wells, in the north east part of the study area (Fig. 11). This comparison reduces subjectivity of the methodology and makes it very promising.

In order to apply the above described methodology all factors were introduced as map layers in the GIS program. Raster data were processed with Vertical Mapper ver. 3.1 [5], whereas Fuzzy Factor Standardization and the multicriteria evaluation has been performed using the GIS program Idrisi Kilimanjaro [3].

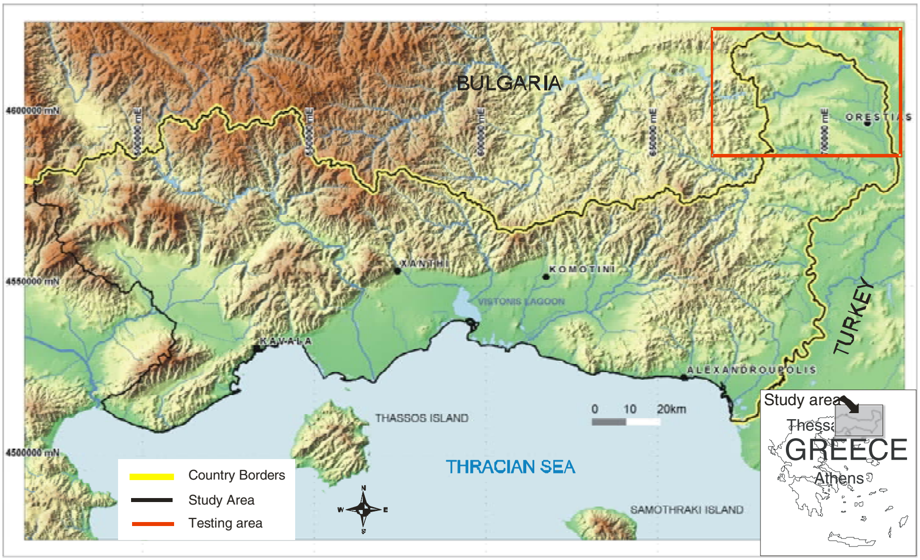


Fig. 11. Location map of the study area

4.1 Methodology Application

Factor Group 1 – Assignment of intrinsic aquifer vulnerability
As mentioned earlier, factor group 1 comprises factors related to the intrinsic aquifer characteristics, i.e., aquifer type, hydraulic conductivity and depth to water table (Fig. 12).

Factor 1 - Aquifer type: Four discrete categories of aquifer types were assigned in the study area, i.e., unconfined, confined aquifer, karst (limestone), and fractured aquifers (groundwater in igneous and metamorphic rocks) (Fig. 12a). While most factors can be automatically rescaled using some mathematical function, rescaling categorical data such as aquifer types requires giving a rating to each category based on some knowledge, according to their relative vulnerability to pollution. In this case, the aquifer vulnerability rating is specified assuming that the most vulnerable ones are karst aquifers which received the highest rating value, i.e., 255. Unconfined aquifers are the next most vulnerable which received a value of 155. Confined aquifers were assigned a value of 75, whereas fractured aquifers are treated as the less vulnerable ones receiving a value of 0. Karst aquifers comprise karstified limestones and marbles of various ages, from the middle Mesozoic to the Eocene. Karstification results in high hydraulic conductivity values and consequently vulnerability to pollution is particularly high. Karst aquifers occupy 13.3% of the study area (1808 km^2). Unconfined aquifers are pore aquifers including mainly the most recent alluvial deposits of the Quaternary to the upper Miocene age. As they are not protected by a confining layer, any potential pollutant released on the ground surface may easily reach groundwater.

Thus, unconfined aquifers are considered as the second most vulnerable to pollution. They form the 25.2% of the study area, i.e., 3427 km^2. Confined aquifers are protected by a confining layer, so pollutants are not expected to reach easily groundwater. They comprise mainly of sedimentary and metasedimentary formations of the late Mesozoic to the Pliocene age, with an aerial extent of 1863 km^2 (13.7% of the study area). The least vulnerable aquifers are the fractured ones, which cannot be considered as ordinary aquifers, in the sense that they are not water bearing formations. They only carry water through fractures and faults due to tectonic events. In the study area, geological formations that include fractured aquifers occupy 47.8% of the study area (6500 km^2), located mainly in the northern mountainous part and are considered as the old rock formations, forming part of the Hellenic hinterland of the Palaeozoic or even older age. The rock types forming these faulted aquifers are gneiss, amphibolites, leptinites, granodiorites, marbles, migmatites, metabasites, ultrabasites. As they have been affected by the pre-alpine tectonics, as well as by the alpine deformation during tertiary times, they appear to be faulted and fractured.

Factor 2 - Hydraulic conductivity: A well inventory has been created including over 2000 boreholes in the study area. Data has been collected from a variety of sources, such as the Greek Geological Survey [49], [50], [51[, [52] and Democritus University of Thrace, as well as from individual drillers and various previous studies [53], [54], [55], [56], [57]. In almost 50 boreholes, pumping tests were performed and the hydraulic properties of the corresponding aquifers were calculated. In areas where no data exist, hydraulic conductivity values were assigned based on bibliographical evidence [58], type of geological formation, grain size and degree of consolidation. Typical hydraulic conductivity values are: 10^{-7} to 10^{-9} m/s for metamorphic rocks (i.e., gneiss, schists, migmatites, metabasites, amphibolites, non-karstified marbles), 10^{-8} to 10^{-7} m/s for clays and marls, 10^{-6} m/s for flysch formations, 10^{-5} to 10^{-3} m/s for sands according to their fine grain content, 10-2 m/s to 10-3 m/s for gravels and conglomerates, 10^{-2} to 10^{-1} m/s to karstified marbles and limestones (Fig. 12b). A monotonically decreasing sigmoidal fuzzy membership function was applied in order to transform the above hydraulic conductivity values to a continuous set of values ranging from 0 to 255. Control points were: 10^{-1} where function membership becomes 1 (highest groundwater vulnerability) and 10^{-9} where function membership becomes 0 (lowest groundwater vulnerability).

Factor 3 - Depth to water table: Depth to water table has been assigned using the well inventory discussed in the previous section. Moreover, data from 168 springs were incorporated in the well inventory. The spring locations were assigned a zero depth to water table, whereas a negative value of depth to water table, i.e., -1 indicates areas where artesian aquifers overflow. A 500-meter buffer zone was created around each borehole and spring, and the depth to water table for the whole zone was assigned as equal to the depth to water table of the associated point. In fractured aquifers the depth to water table was assigned an arbitrary value of 100, as no groundwater level is present in these rock formations, and consequently, a relatively high value for depth to water table is needed for calculations to proceed (Fig. 12c).

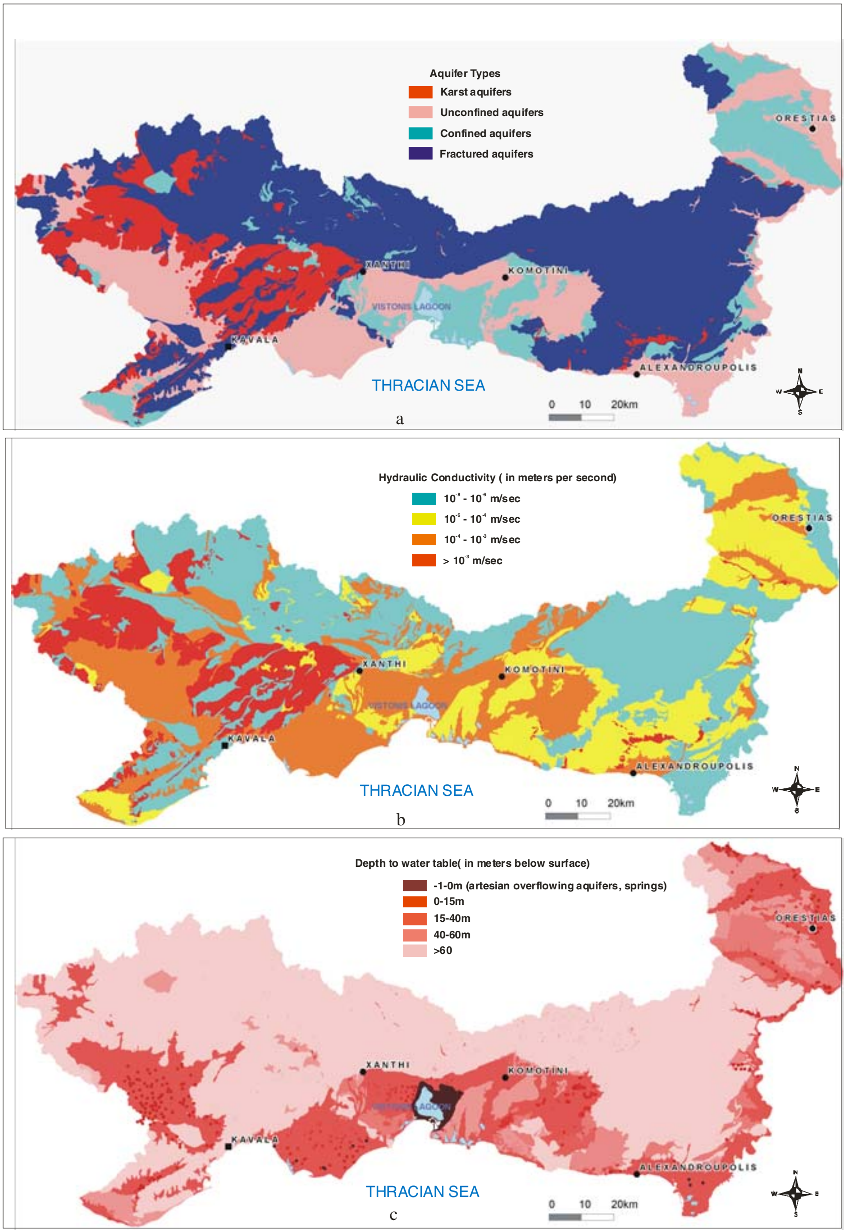


Fig. 12. Factors of the internal aquifer system properties.

A monotonically decreasing sigmoidal fuzzy membership function was applied in order to transform the above depth to water table values to a continuous set of values ranging from 0 to 255. Control points were: 0 where function membership becomes 1 (highest groundwater vulnerability) and 100 where function membership becomes 0 (lowest groundwater vulnerability). Values greater than 100 receive the same function membership value of 0, as higher values of depth to water table were considered to have no more impact on groundwater vulnerability. In areas where overflowing occurs (artesian aquifers with piezometric surface above topographic surface) an arbitrary value of 25 (low vulnerability) in the scale of 0 to 255 has been assigned, as in this particular case the water is flowing out of the aquifer, minimizing thus the potential of groundwater pollution in case a pollutant is released.

Factor Group 2 – Assignment of aquifer vulnerability related to external forces: Factor group 2 comprises external factors to the aquifer system, i.e., surface runoff, non-concentrated land uses, proximity to concentrated land uses, proximity to major rivers, proximity to residential areas, areas protected by national law or international environmental treaties, and proximity to highways and railways (Fig. 13, 14, 15, 16).

Factor 4 - Surface runoff: Each pixel in the study area receives an amount of rain (Fig. 13a). RUNOFF is an especially developed routine incorporated in the GIS program Idrisi Kilimanjaro, that calculates the accumulation of rainfall units per pixel based on an elevation image (DEM)[3] (Fig. 13b), and is a modification of the algorithm described by Jenson and Domingue [32]. The higher the accumulation of runoff in a pixel the most vulnerable to groundwater pollution the pixel is. A simple RUNOFF analysis accumulates rainfall on a per pixel basis as if one unit of rainfall was dropped on every location. In our case, mean annual precipitation data from 82 gaging stations scattered all over the study area, for the time period 1966 – 2001, was incorporated in the routine. Besides providing a DEM, the routine has been enhanced with a permeability image (Fig. 13c), in order to account for soil permeability and to adjust the accumulated rainfall accordingly (Fig. 13d) [3].

A monotonically increasing sigmoidal fuzzy membership function was applied to transform runoff accumulation to a 0 to 255 scale. Control points were: 600 where function membership becomes 0, and 35,000 mm/yr where function membership becomes 1. The value of 35,000 mm/yr after multiplication by the pixel size, i.e., 3600 m^2, corresponds to an average discharge of 345 m^3/d. The first and the second control points were selected after imposing a quantiles classification scheme, placing an equal number of pixels into each class. The end point of the first class, i.e., 600 mm/yr (approx. 5% of the study area) and the starting point of the last class, i.e., 35,000 mm/yr (basin outlets) (upper 5% of the study area) defined as control points 1 and 2, respectively, for the sigmoidal fuzzy function. Pixels with accumulated runoff values lower than 600 mm/yr and higher than 35,000 mm/yr were considered equally vulnerable to the first and second control points and received a function membership value of 0 (lowest vulnerability) and 1 (highest vulnerability), respectively.

Factor 5 - Non-concentrated land uses: This factor includes all types of land uses that, unlike industrial land uses, cover a wide range of the study area. The following 6 discrete land use categories are introduced in the calculation process: Forest dense, Forest sparse, Grasslands, Water bodies, Agricultural and Urban (Fig. 14a). As in the case of aquifer type factor, rescaling categorical data such as land uses requires giving a rating to each category based on some knowledge [3], according to their relative groundwater vulnerability. On the continuous 0 to 255 scale, a vulnerability rating of 255 has been assigned to agricultural and urban land uses as they are considered to be equally polluting land uses, worse than any other non-concentrated human activity, as far as groundwater pollution is concerned. A rating of 150 was assigned to grasslands, where most of the livestock farming takes place. A value of 30 has been assigned to sparse forest and 0 to dense forest and to water bodies (least groundwater vulnerable areas as far as land uses are concerned).

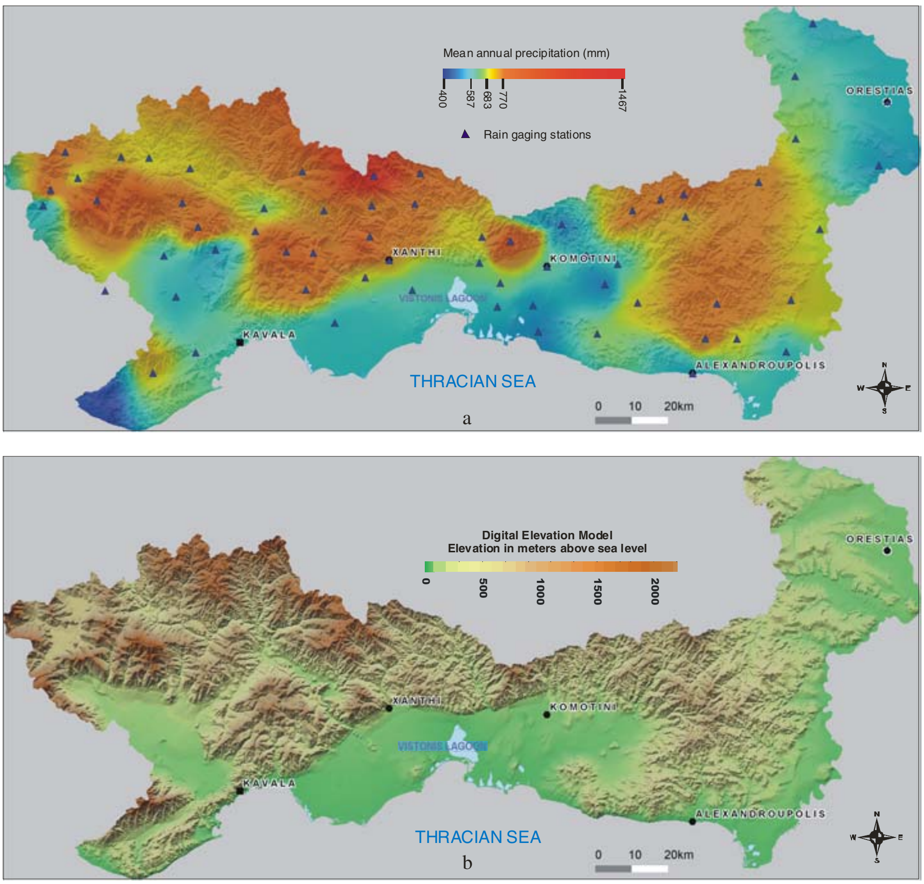


Fig. 13. (a) Precipitation distribution in the study area; (b) Digital Elevation Model; (c) Permeability of geological formations; (d) Surface runoff accumulation

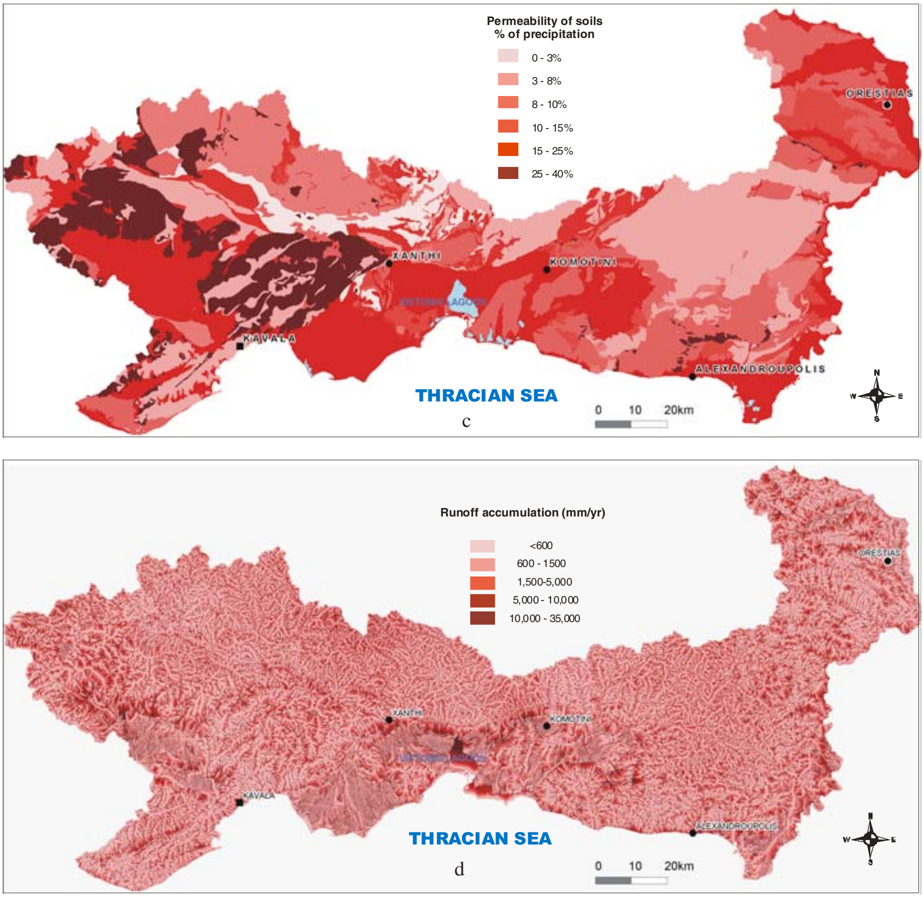


Fig. 13. (*continued*)

Factor 6 - Proximity to concentrated land uses: Concentrated land uses include human activities that may result in groundwater pollution, if no protective measures are taken. The following major concentrated human activities are distinguished in the study area: landfills, industrial areas, airports, oil tanks, ports, salt works and pumping stations for irrigation returns (Fig. 14a). Facilities such as gas stations, septic tanks and wastewater treatment units are considered to be part of the urban area and are not examined under this factor. All the above activities are not supposed to have the same degree of vulnerability, i.e., landfill leachates are far more polluting than pumping stations for irrigation returns. An importance weight for each of the above activities has been assigned as follows: landfills 1000, industrial areas 750, ports, airports, oil tanks, salt works and irrigation return pumping stations 500. A location profile has been created, representing the average distance to a series of points from anywhere within a map area [5] (Fig. 14b). The algorithm generates a grid, where at each cell a value is calculated that represents the weighted average distance to all point locations surrounding that cell, according to the relative importance weight of each point lying

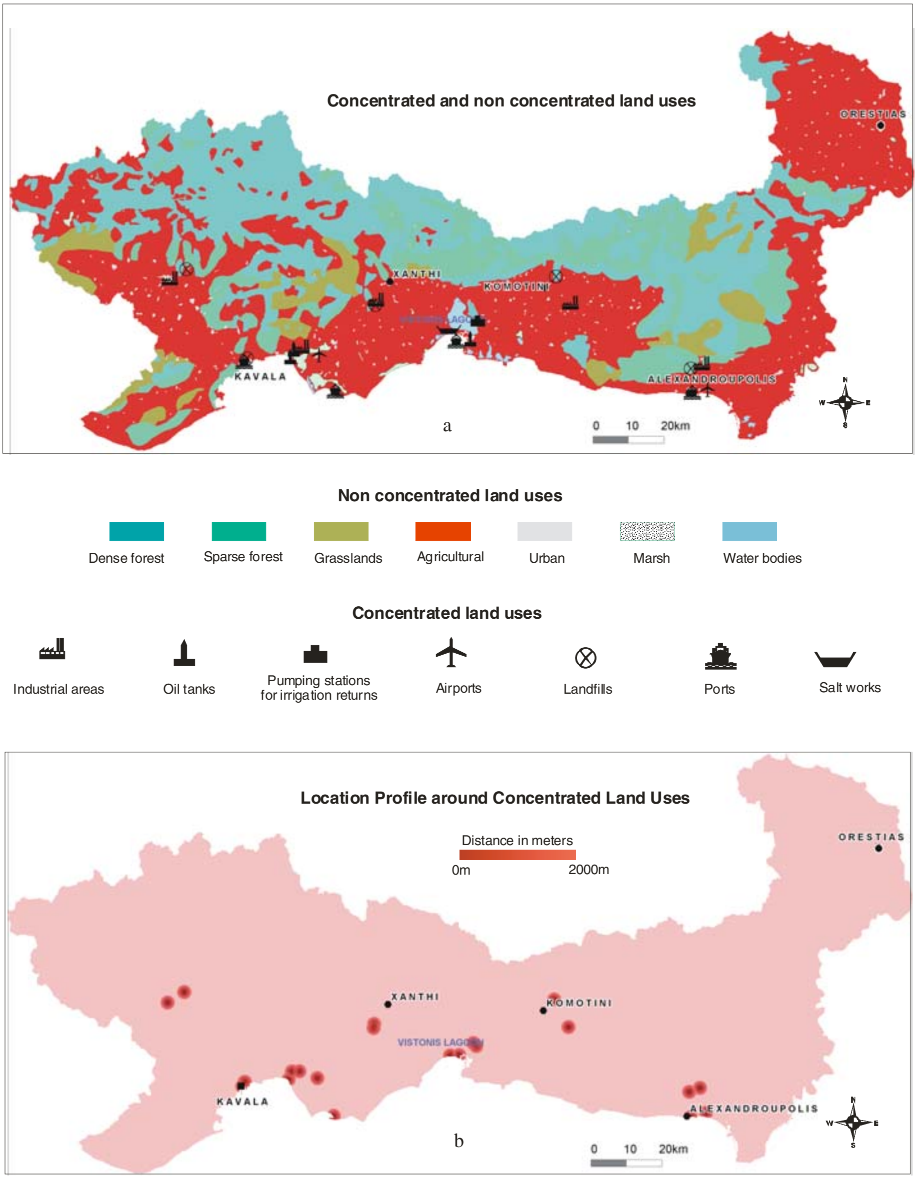


Fig. 14. (a) Concentrated and non-concentrated land uses; (b) Location profile around concentrated activities

within a defined search radius. In the present study, a search radius of 2000 meters was applied, as it corresponds to approximately a 5-year travel time distance, taking an average groundwater velocity of approximately 1.1m/day [59], and it is supposed to be the maximum radius of influence for all the above mentioned activities.

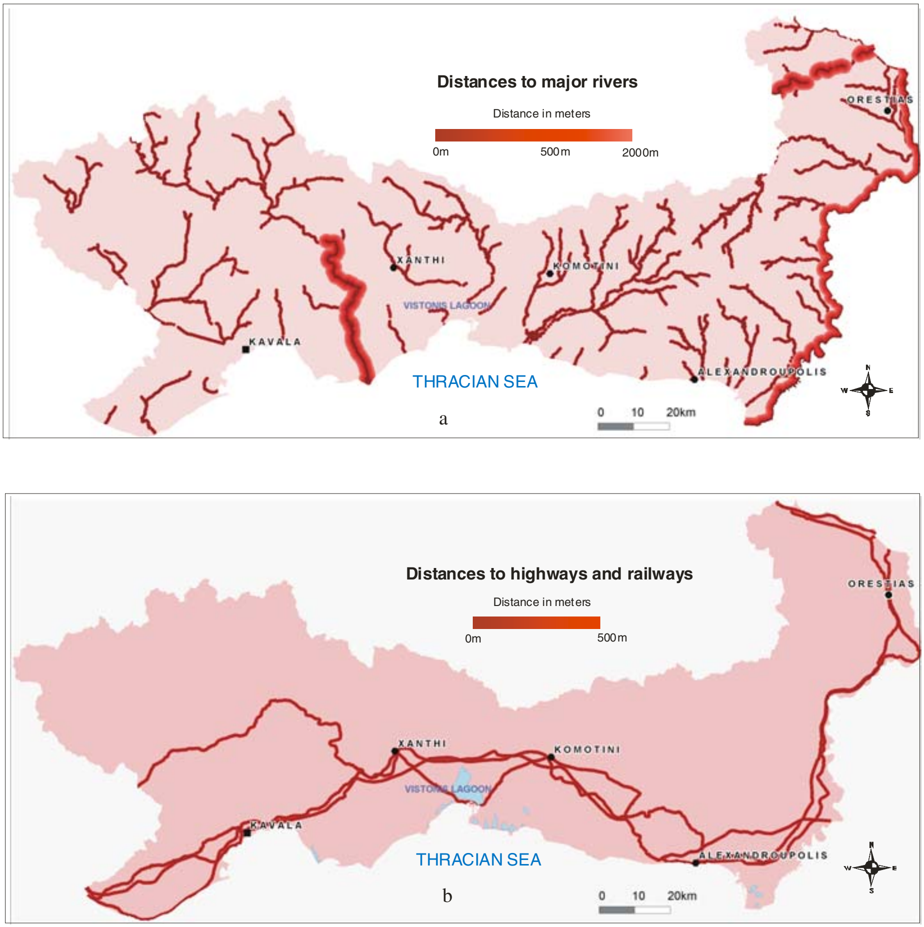


Fig. 15. Proximity to (a) major rivers; (b) highways and railway in the study area

A monotonically decreasing sigmoidal fuzzy membership function was applied in order to transform the calculated distances to a continuous set of values ranging from 0 to 255. Control points were: 0 where function membership becomes 1 and 2000 where function membership becomes 0. Distances greater than 2000 m from concentrated activities are considered to have the same membership function value of 0, as points at those distances are not expected to be vulnerable to groundwater pollution due to these activities.

Factor 7 - Proximity to major rivers: Each river is a potential final receiver of treated or even untreated wastewater, so aquifers close to river beds, and especially those that are hydraulically connected to them, are expected to receive part of the pollutants of wastewaters. In the present study distances to major rivers were calculated and were transformed to a 0 to 255 scale using a monotonically decreasing sigmoidal fuzzy membership function. Different control points were used for the main three rivers of the study area (Nestos, Evros and Ardas), which are evidently hydraulically connected to the underlying aquifers [49], [52], and for those rivers of less importance, where there is no evidence for hydraulic connection to adjacent aquifers. In that way, control points for the first river category were: 0 m where function membership becomes 1, and 1000 m where function membership becomes 0. Distances greater than 1000 m from river beds are considered to have the same function membership value of 0. Similarly, control points for the second river category were: 0 m where function membership becomes 1, and 500 m where function membership becomes 0 (Fig. 15a).

Factor 8 - Proximity to residential areas: While analyzing the non-concentrated land use factor, a land use category has been the urban one. Besides the presence of an urban center, groundwater vulnerability to pollution is also influenced by the number of inhabitants, as well as by the distance to any potential aquifer. These are two parameters that were not incorporated in the land use factor. A zone of influence around each city, town and human settlement has been created, with radius proportional to their inhabitants, according to the 2001 census data provided by the National Statistical Service of Greece (Fig. 16a). Each zone of influence has been assigned the value of the town population, thus urban areas with higher population are considered to have influence in greater distances and to a higher extent.

A monotonically increasing sigmoidal fuzzy membership function was used to convert data to a 0 to 255 scale, i.e., the vulnerability of groundwater to pollution increases proportionally to population. Control points were set to 56,000 (the city with the highest population in the study area) where the fuzzy membership function becomes 1, and 0 where fuzzy membership function becomes 0, i.e., groundwater vulnerability is diminished in areas with no population present.

Factor 9 - Areas protected by national law or international environmental treaties: The study area contains wetlands of international interest, such as the Delta of River Evros, the Delta of River Nestos, and the lakes Vistonida and Ismarida. Several areas belong to the European network Natura 2000 or to National Parks [60], [21] and should be protected according to the Greek law, EU conservation policies or international treaties, such as the Ramsar Convention of 1971 (Fig. 16b). Most industrial and agricultural activities are prohibited in these areas, thus the presence of such protected regions can be considered as a protective factor to groundwater pollution. In this case, two distinct categories were assigned: the first category includes all areas that belong to National Parks, Natura 2000 sites, National Forests and other protected regions and received a value of 0 (least vulnerability to pollution), whereas all other areas received a value of 255 for this particular factor.

Factor 10 - Proximity to highways and railways: Highways and train railways are the main routes of transport for harmful chemicals and are potential sources of groundwater pollution, in case of an accident. Thus, a zone of influence of 500 m was created for highways and railways of the study area (Fig. 15b). A monotonically decreasing sigmoidal fuzzy membership was applied to transform distances to 0 to 255 scale. Control points were set to 0 (highest vulnerability) where the membership function equals 1, and 500 (the distance of influence in case an accident occurs) where the fuzzy membership function becomes 0.

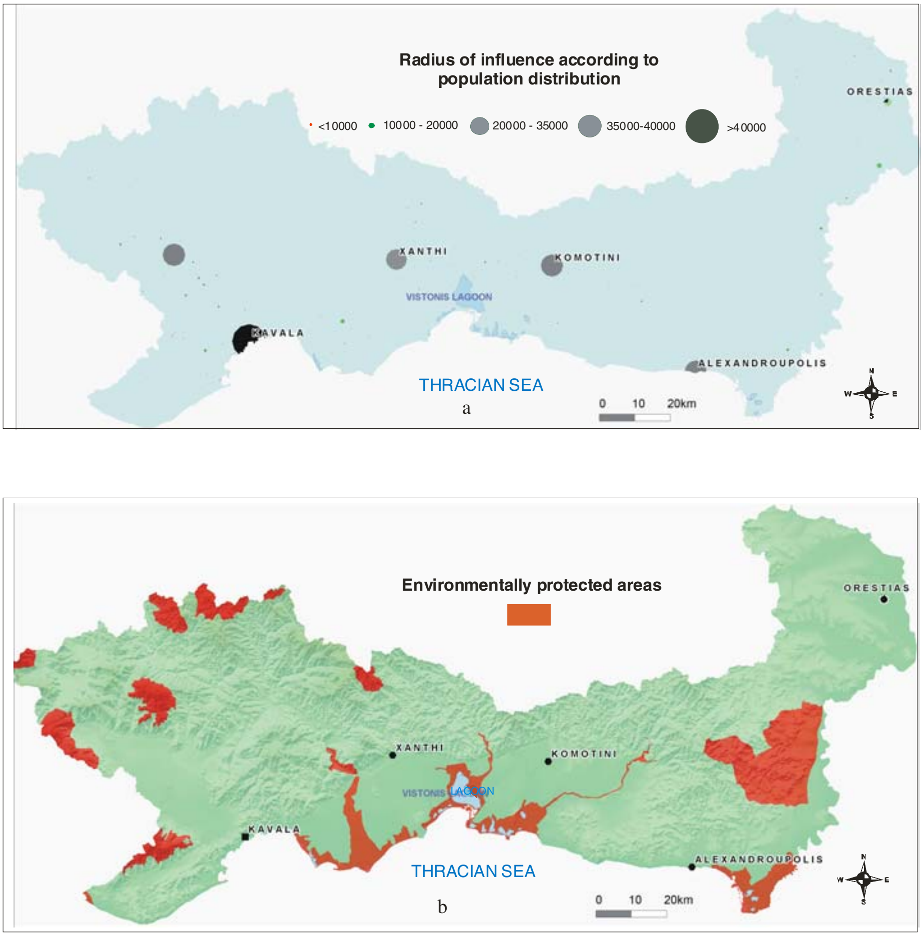
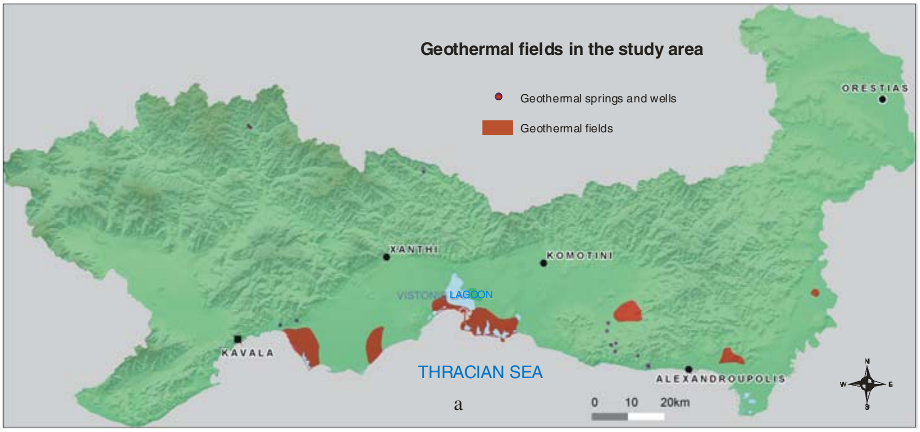


Fig. 16. (a) Radius of influence of residential areas according to population distribution; (b) environmentally protected areas

Factor Group 3 – Assignment of aquifer vulnerability related the presence of local geological conditions.

Factor group 3 comprises two factors related to the presence of particular geological conditions that influence locally the aquifer system. These factors are the presence of geothermal fields and the presence of salt water intrusion zones, which are two common features in the study area (Fig. 17).



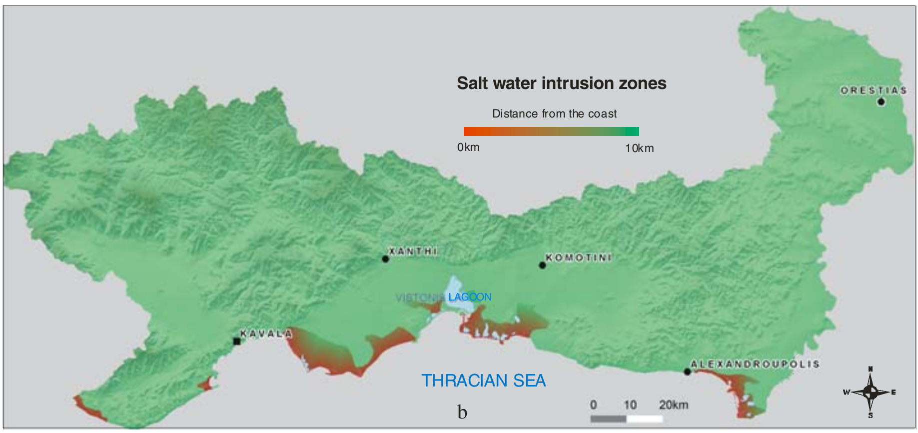


Fig. 17. Factors of the local geological conditions: (a) presence of a geothermal field; b) presence of a saltwater intrusion zone

Factor 11 - Presence of geothermal fields: Several geothermal fields are present in the study area, influencing both the temperature and chemistry of groundwater (Fig. 17a). Known geothermal fields were delineated and received a value of 255, whereas a zone of influence of 500 m around each known geothermal borehole or spring has been created. A monotonically decreasing sigmoidal fuzzy membership function was applied in order to transform the distances away from geothermal boreholes or springs to continuous set of values ranging from 0 to 255. Control points were set to 0, where function membership becomes 1 (highest vulnerability), and 500, where function membership becomes 0 and groundwater vulnerability is diminished and becomes independent of distance.

Factor 12 - Presence of salt water intrusion zones: In areas close to the coast and where geological conditions are favorable, salt water intrusion has been detected at distances even 10 km away from the coast. Areas of salt intrusion were delineated in the study area, according to borehole data [53] (Fig. 17b). However, the influence of salt intrusion is considered to diminish proportionally to the distance from the coast; thus a monotonically decreasing sigmoidal fuzzy membership function was applied, with control points set to 0 (coastline, highest groundwater vulnerability) where the function becomes 1, to 10 km away from the coast (lowest groundwater vulnerability), where function membership becomes 0 and groundwater vulnerability is diminished and becomes independent of distance from the coast.

4.2 Aggregation Procedure and Results

After standardizing all factors to a common 0 to 255 scale using fuzzy membership functions, factor weights were given to all factors in each group. The weights indicate a factor's importance relative to all other factors and they control how factors compensate for each other. In other words, the degree to which a factor compensates for another is determined through its factor weight. In the case of determining groundwater vulnerability to pollution factors with low groundwater vulnerability to pollution in a given location can compensate for other factors with high groundwater vulnerability in the same location.

In order to combine information from various factors in each factor group, an aggregation procedure should be applied. The most commonly used vulnerability assessment methods, like DRASTIC, apply an aggregation technique known as Weighted Linear Combination [1], where factor scores are multiplied by their factor weight and then summed to yield the vulnerability score:

$$S = \sum_{i=1}^{N} w_i x_i \cdot \tag{5}$$

where: S = vulnerability score; w_i = weight of factor i; x_i = criterion score of factor i; N = number of factors.

However, when applying the same factor weights in an extended area, involves the risk of underestimating the local particularities, i.e., salt water intrusion may be a factor of low or no importance for most of the study area, but a crucial factor for groundwater vulnerability in the coastal zones. Honoring the local characteristics of the system, while examining data at the regional scale, is achieved by distinction of factors into three factor groups and by applying a second set of weights in each factor group, known as order weights. This second set of weights controls the manner in which the weighted factors are aggregated [34], [35].

After factor weights are applied to the original factors, the results are ranked from low to high groundwater vulnerability to pollution for each location. The factor with the lowest vulnerability score is then given the first order weight, the factor with the next higher vulnerability score is given the second order weight, and so on. This has the effect of weighting factors based on their rank from minimum (lowest vulnerability) to maximum (highest vulnerability) value for each location. The relative skew toward either minimum or maximum of the order weights controls the level of risk in the evaluation [3]. Additionally, the degree to which the order weights are evenly distributed across all positions controls the degree to which factor weights have influence. The procedure is repeated once for each group of factors, resulting in three intermediate vulnerability maps. The final vulnerability map is produced by aggregating, the intermediate results using the same procedure. Variables are rated according to the 9-point scale shown in Table 2.

Weighting and aggregating factor group 1: The factor weights assigned in factor group 1 are shown in Table 6.

Table 6. Pairwise comparison matrix for factor group 1

	Aquifer type	Hydraulic conductivity	Depth to water	Calculated factor weights
Aquifer type	1	5	3	0.6370
Hydraulic conductivity	1/5	1	1/3	0.1047
Depth to water	1/3	3	1	0.2583

Consistency ratio = 0.03 (< 0.1 acceptable)

Order Weights:	0.2	0.3	0.5
Rank:	1st	2nd	3rd

The assigned order weights show a moderate level of risk. Skew towards the factor with the higher vulnerability score, shows a strict decision. The intermediate intrinsic aquifer vulnerability map is presented in Figure 18a.

Weighting and aggregating factor group 2: The factor weights assigned in factor group 2 are shown in Table 7.

Table 7. Pairwise comparison matrix for factor group 2

	F1	F2	F3	F4	F5	F6	F7	Calculated weights
F1	1	3	5	3	5	3	7	0.3606
F2	1/3	1	3	1	3	1	5	0.1541
F3	1/5	1/3	1	1/3	1/3	1/3	3	0.0546
F4	1/3	1	3	1	5	1	5	0.1739
F5	1/5	1/3	3	1/5	1	1/3	3	0.0728
F6	1/3	1	3	1	3	1	5	0.1541
F7	1/7	1/5	1/3	1/5	1/3	1/5	1	0.0299

F1: Surface runoff accumulation; F2: Non-concentrated land uses; F3: Areas protected by national law or international environmental treaties; F4: Proximity to concentrated land uses; F5: Proximity to residential areas; F6: Proximity to major rivers; F7: Proximity to highways and railways;

Consistency ratio = 0.05 (acceptable)

Order Weights:	0.1	0.1	0.1	0.1	0.15	0.2	0.25
Rank:	1st	2nd	3rd	4th	5th	6th	7th

The assigned order weights show a moderate level of risk and a strict decision (the three factors with the highest score share 50% of the weight, while all the rest are equally weighted). Figure 18b presents the groundwater vulnerability to pollution based on factor group 2.

Weighting and aggregating factor group 3: The following factor weights were assigned in factor group 3:

Presence of geothermal fields: 0.5; Presence of salt water intrusion zones: 0.5 (factors are equally weighted).

The following order weights were assigned in factor group 3:

Order Weights:	0.4	0.6
Rank:	1st	2nd

The assigned order weights show a moderate level of risk. Skew towards the factor with the higher vulnerability score shows a strict decision. Figure 18c shows the groundwater vulnerability based on factor group 3.

Weighting and aggregating intermediate results: Intermediate results were aggregated in order to produce the final groundwater vulnerability map. Two aggregation procedures were performed resulting in two groundwater vulnerability to pollution maps for the study area. The intermediate vulnerability maps, which resulted from the three previously examined groups, were used as factors in these final aggregation processes. Groundwater vulnerability results were grouped in 5 distinct classes as shown in Table 10 (columns 3 and 4) and Figures 18d and 19.

A first aggregation procedure, without fitting the calculated values to observed ones, was attempted, using equal weights for factor groups 1 and 2, and skewing order weights towards the factor group that shows the greatest vulnerability score. In that way, the factor weights and order weights used in this first aggregation are shown in Table 8.

Table 8. Pairwise comparison matrix for the final aggregation procedure

	Factor group 1	Factor group 2	Factor group 3	Calculated factor weights
Factor group 1	1	1	3	0.4286
Factor group 2	1	1	3	0.4286
Factor group 3	1/3	1/3	1	0.1429
Consistency ratio =	0.00	(best)		
Order Weights:	0.2	0.3	0.5	
Rank:	1st	2nd	3rd	

The produced vulnerability map is presented in Figure 18d and in percentages in column 3 of Table 10.

In the second aggregation, an attempt was made to calibrate the aggregation procedure by comparing with measured groundwater nitrate concentration data from 141 boreholes in the NE part of the study area, which consists of alluvial aquifers [52]. Rather than examining the accurate estimations of nitrate concentrations, the comparison focused on the likelihood of a location being classified as contaminated [48]. The categories were chosen to indicate low anthropogenic effect (<0.5 mg/L), moderate effect (0.5 - 3 mg/L), moderately high effect (3 - 10 mg/L), high effect (10 - 50 mg/L) and very high anthropogenic effect (>50 mg/L) (Fig. 19). Results show best fit when factor group 1, i.e., factors related to the hydraulic parameters of the internal aquifer system, takes most of the weight. The factor weights assigned in the final aggregation are shown in Table 9.

Table 9. Pairwise comparison matrix for the final aggregation procedure after calibration with measured data

	Factor group 1	Factor group 2	Factor group 3	Calculated factor weights
Factor group 1	1	5	7	0.7306
Factor group 2	1/5	1	3	0.1884
Factor group 3	1/7	1/3	1	0.0810
Consistency ratio = 0.06 (< 0.1 acceptable)				
Order Weights:	0.1	0.4	0.5	
Rank:	1st	2nd	3rd	

The assigned order weights show a low level of risk. Aggregation is clearly skewed towards the factor with the highest vulnerability score.

Table 10. Distinct groundwater vulnerability categories in the study area

Groundwater Vulnerability Index	Vulnerability Class	Percentage of the study area falling in the groundwater vulnerability class	
		before calibration with measured data	after calibration with measured data
0 -50	Low	22%	11%
50 -100	Moderate	28%	35%
100 – 150	Moderately high	34%	31%
150 - 200	High	15%	21%
200 - 255	Very high	1%	2%

A coincidence report was generated between well contamination data and the composite groundwater vulnerability map created after integrating GIS, fuzzy logic and multicriteria evaluation methods (Fig. 19). In this report it is demonstrated that 82 wells from the 141 tested wells were categorized in the correct groundwater vulnerability class (those located on the diagonal of Table 11); 4 wells were categorized in one vulnerability class higher than the observed class (those located below the diagonal of Table 11); 41 wells were categorized in one vulnerability class lower than the observed (those located above the diagonal of Table 11); 11 wells were classified in two vulnerability classes lower and 3 wells were categorized in three classes lower.

Table 11. Coincidence of wells with five contamination levels and vulnerability classes as estimated by the presented methodology

	Concentration of nitrate in groundwater (mg/L)				
Model	Low (< 0.5)	Moderate (0.5 - 3)	Moderately high (3 - 10)	High (10 - 50)	Very high (> 50)
Low	13	6	4	3	
Moderate	3	11	10	2	
Moderately high		1	20	23	5
High				36	2
Very high					2

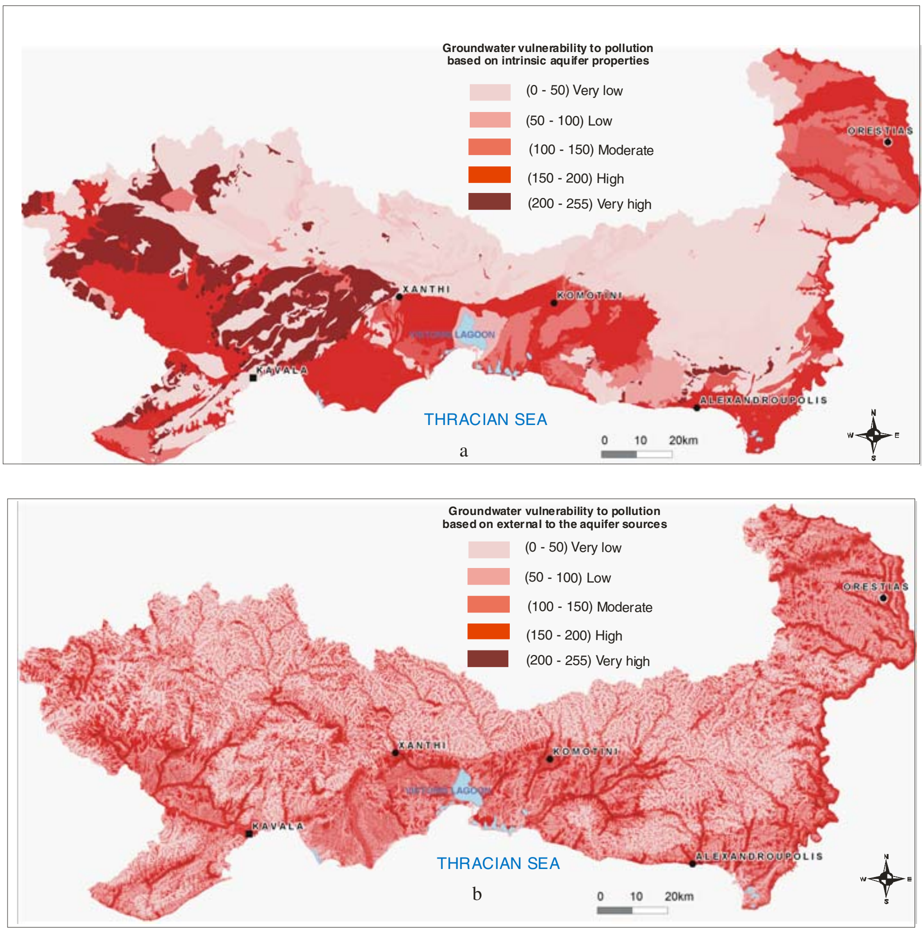
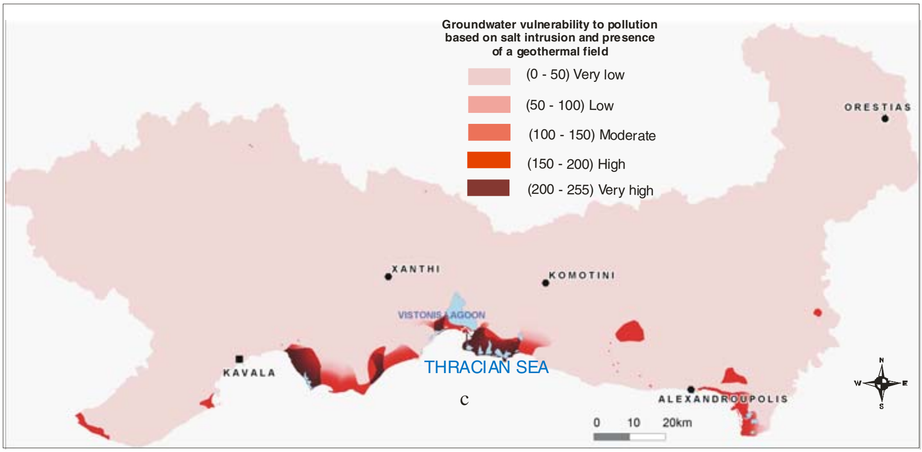


Fig. 18. Intermediate and composite groundwater vulnerability maps: (a) produced analyzing factor group 1; (b) produced analyzing factor group 2; (c) produced analyzing factor group 3; (d) composite groundwater vulnerability map produced aggregating the three intermediate vulnerability maps



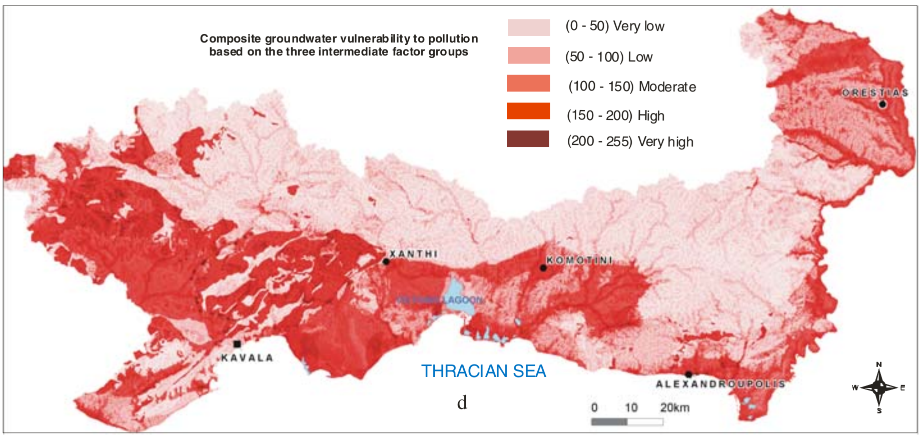


Fig. 18. (*continued*)

4.3 Discussion

In this study of assessing groundwater vulnerability to pollution, a critical issue is the standardization of factors and specifically the selection of function control points (inflection points) of the fuzzy membership function, which is based on the understanding of the criterion. As far as distances are concerned, inflection points are located at zero distance, up to the distance where the effect of activity under consideration on groundwater is diminished. In case of factors of broad value range, such as runoff accumulation, a quantiles classification scheme was applied, placing an equal number of pixels into each class, and control points were located at the end of the first class and the starting point of the last class (approximately upper and lower 5% of the study area).

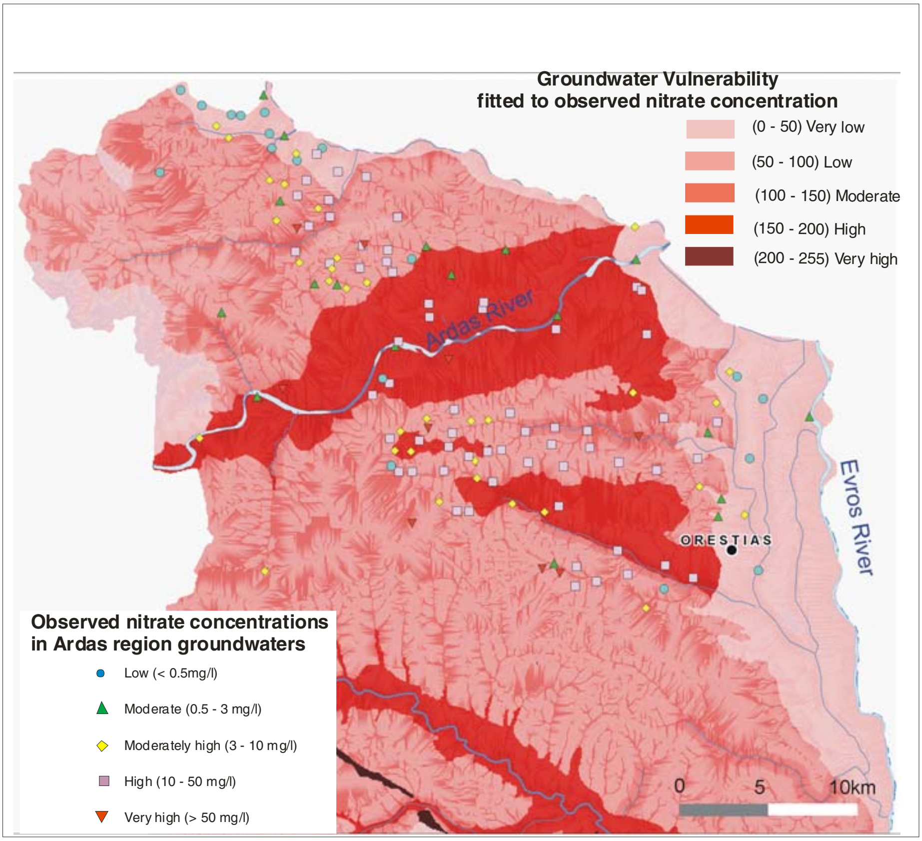


Fig. 19. Composite groundwater vulnerability map fitted to observed nitrate concentrations in groundwater

Factor weights were assigned to the twelve factors involved in the calculation process. It is clear that assignment of factor weights is based on previous knowledge of the aquifer characteristics and the particularities of the study area, as well as on the experience of the scientists involved in the weight assignment process. It was attempted, however, to develop a weight assigning process as objective as possible, applying techniques like the Analytical Hierarchy Process and the Order Weighted Averaging. A best fit to the measured nitrate concentration data, in a sub-region of the study area, is achieved only when factors related to the intrinsic aquifer characteristics were assigned most of the weight. This is due to the fact that in the tested area agricultural land uses prevail, polluting the associated aquifers with related pollutants. The main factor that differentiates contamination throughout this region is hydraulic conductivity, which is particularly low in the eastern part of this region, thus leading to lower nitrate concentrations in groundwater.

Another interesting aspect of the study is the fact that areas considered to be the most vulnerable ones, especially karst aquifers in the western part of the study area, do not demonstrate contamination of their groundwaters. Having in mind that groundwater vulnerability to pollution is a relative, non-measurable property [45] and that it only shows the likelihood or risk for contaminants to reach the groundwater system after introduction at some location above the uppermost aquifer [61], [62]. Therefore, one should be careful when trying a comparison between the observed and the estimated contamination of groundwaters, especially in areas with contradictory properties. In addition, trying to calibrate parameters in the calculation process, such as factor weights, might lead to serious misinterpretations, especially in areas with high estimated groundwater vulnerability, where little or no contamination has been detected.

The presented work, in both aggregations used, categorizes almost 60% of the study area to moderate and moderately high vulnerability classes. From this point of view the presented methodology is closer to the DRASTIC method [37] based on the comparison presented by Gogu et al. [46]. Nevertheless, DRASTIC and EPIK methods classify limestone aquifers to moderate vulnerability [46], whereas the presented methodology assess them to high vulnerability like the German method [43].

Besides the differences that all the above mentioned methods present, it is true that the choice of the vulnerability method is still a subjective decision which is restricted by the available information concerning aquifer systems. Many problems and inefficiencies are expected to be overcome when the groundwater vulnerability assignment process is coupled with groundwater flow and transport modeling and the results are subject to correlation with observed contamination data. It is true, though, that coupling with groundwater modeling techniques is almost impossible at the regional scale, especially when karst aquifers are involved. However, results provided by groundwater vulnerability to pollution estimation at the regional scale, with the methodology outlined in the present study, are very useful for future siting of human activities, urban development and water resources management, as well as for highlighting areas of particular interest, where coupling with groundwater modeling, at a finer scale, could offer a better understanding of the aquifer system.

5 Conclusions

Population increase and growth perspectives in the living standards have resulted during recent years in pressures on the environment. Resource allocation and exploitation techniques are becoming more intense and consequently considerable interest is focused on the use of GIS as a decision support tool for natural resources management. The presented herein methodology shows how the GIS technology in combination with other modern tools such as Multi Criteria Decision Making and Fuzzy Logic can be applied while dealing with several environmental problems. The proposed approach uses either exclusionary and/or non–exclusionary criteria in the decision making process. The application of exclusionary criteria only, as presented in the application of section 2, provides a fast way of evaluating possible sites, however it offers no flexibility in the risk and the trade off of the final decision. It is very convenient for an initial screening of extended areas. Application of exclusionary and non-exclusionary criteria is presented in section 2, in a problem of siting a MSW landfill. In this case, the AHP was used and offered a quite objective weights assignment process. Furthermore, the use of the second set of weights, order weights, as applied in the present work, provides great flexibility in the aggregation procedure, offering full control over the risk and trade off in the final decision. The third application, presented in section 4, applies only non-exclusionary criteria. The aim, this time, is not to site a specified activity, but to assign a groundwater vulnerability value to each pixel of the area of interest. Thus, no pixels are excluded from the computational process and only non-exclusionary criteria are applied. The methodology is quite the same as the one presented in section 3 and may well be applied for evaluating other environmental properties such as flood risk, landslide susceptibility or seismic risk among others. Nevertheless, it should be pointed out here, that the combination of the GIS technology with the AHP and the Fuzzy Logic forms a promising and powerful tool, but an equally important ingredient of the decision making process is the ability of the decision maker to select and combine in the most appropriate way the several criteria, depending on the nature of the objective. For example a stated objective might have as perspective to minimize the effects of timber harvesting on recreational uses in the area. The criteria and the weights to be assigned would be quite different from that of another objective that targets mainly on profit maximization. Thus, primary issues always addressed in environmental problems are the motives and the social perspective and usually they change dramatically the outcome of the whole procedure.

Resources

A Key books

Abdul-Rahman. A., Pilouk. M.: Spatial Data Modelling for 3D GIS. Springer-Verlag Berlin Heidelberg (2008)

Bourrough, P.A.: Principles in Geographical Information Systems for Land Resources Assessment. 8th edn. Clarendon Press, Oxford (1992)

Brewer, C.A.: Designing Better Maps: A Guide for GIS Users. ESRI Press, Redlands, California (2005)

Demicco, R.V., Klir, G.J.: (eds.): Fuzzy Logic in Geology. Elsevier Academic Press (2003)

Eastman, J.R.: IDRISI Kilimanjaro: Guide to GIS and Image Processing. Clark Laboratories, Clark University, Worcester, USA. (2003)

Fetter, C.V.: Applied Hydrogeology. 3rd Edition, Macmillan College Publishing Company (1994)

MapInfo Corp.: Mapinfo Professional ver. 8.5 User Guide. Mapinfo Corp. (2006)

Northwood Technologies Inc. and Malconi Mobile Ltd.: Vertical Mapper Spatial Analysis and Display Software, Version 3.0, User’s Manual. Northwood Technologies Inc., Canada (2001)

Voogd, H.: Multicriteria Evaluation for Urban and Regional Planning. Pion, London (1983)

B Key Survey/Review Articles

Carver, S.J. (1991): Integrating Multi-Criteria Evaluation with Geographical Information Systems, International Journal of Geographical Information Systems 5:321-339

Diamond, J.C., Wright, J.R. (1988): Design of an Integrated Spatial Information System for Multiobjective Land – Use Planning, Environment and Planning B: Planning and Design 15: 205 - 214

Eastman, J.R., Jin, W., Kyem, P.A.K., Toledano, J. (1995): Raster Procedures for Multi-Criteria / Multi Objective Decisions, Photogrammetric Engineering and Remote Sensing 61: 539 - 547

Gogu. R.C., Dassargues. A. (2000): Current trends and future challenges in groundwater vulnerability assessment using overlay and index methods, Environmental Geology 39: 549-559

Saaty, T.L. (1977): A Scaling Method for Priorities in Hierarchical Structures, Journal of Mathematical Psychology 15: 234-281

C Key Journals

Photogrammetric Engineering and Remote Sensing (ASPRS)
Environmental Geology (Springer)
International Journal of Geographical Information Science (Taylor and Francis)
Journal of Environmental Management (Elsevier)
Remote Sensing of Environment (Elsevier)
International Journal of Intelligent Systems (Willey)

D Key International Conferences/Workshops

Fourth European Conference and Exhibition on Geographical Information Systems (EGIS '93)
Second International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences
Expert Group Workshop: Software Technology for Agenda 21: Decision Support Systems for Sustainable Development, The United Nations University, International Institute for Software Technology, February/March 1996
International Conference for the Protection and Restoration of the Environment (PRE)
International Conference on Environmental Management, Engineering, Planning and Economics (CEMEPE)

E Software

The software packages used for our research are the following:
Idrisi Kilimanjaro
http://www.clarklabs.org
MapInfo Professional ver. 8.5
http://www.mapinfo.com
Vertical Mapper ver. 3.1
http://www.mapinfo.com

F Data Bases

All geospatial data used in the above presented case studies is part of the following digital data bases, which were provided by Geoinfo Co. free of charge for the purposes of the present research:
Geoinfo Co., 1998. Digital Data for Greece - GR Survey Data Base.
Geoinfo Co., 2002. Digital Data for Greece - GR Operational Data Base.
http://www.geoingo.gr.

References

1. Voogd, H.: Multicriteria Evaluation for Urban and Regional Planning. Pion, London (1983)
2. Carver, S.J.: Integrating Multi-Criteria Evaluation with Geographical Information Systems. International Journal of Geographical Information Systems 5, 321–339 (1991)
3. Eastman, J.R.: IDRISI Kilimanjaro: Guide to GIS and Image Processing. Clark Laboratories. Clark University, Worcester (2003)
4. Siddiqui, M., Everett, J.M., Vieux, B.E.: Landfill siting using Geographical Information Systems: A demonstration. Journal of Environmental Engineering 122, 515–523 (1996)
5. Northwood Technologies Inc., Malconi Mobile Ltd.: Vertical Mapper Spatial Analysis and Display Software, Version 3.0, User's Manual (2001)
6. Kontos, T.D., Komilis, D.P., Halvadakis, C.P.: Siting MSW landfills on Lesvos island with a GIS-based methodology. Waste Management and Research 21, 262–278 (2003)
7. Geoinfo Co.: Digital Data for Greece - GR Survey Data Base (1998), `http://www.geoingo.gr`
8. Economopoulou, M.A., Tsihrintzis, V.A.: Sensitivity analysis of stabilization pond system design parameters. Environmental Technology 23, 273–286 (2002)
9. Economopoulou, M.A., Tsihrintzis, V.A.: Design methodology of free water surface constructed wetlands. Water Resources Management 18, 541–565 (2004)
10. OASP (Organization of Antiseismic Planning and Protection of Greece): The New Seismic Hazard Map of Greece (2004)
11. Greek Governmental Ministry Decision 114218/97 (1997): Document 1016 B'/17-11-1997. Framework of Specifications and General Planning for the Management of Solid Wastes (in Greek)
12. Dafis, S., Papastergiadou, E., Georgiou, K., Babalonas, D., Georgiadis, T., Papageorgiou, M., Lazaridou, T., Tsiaoussi, V.: Directive 92/43/EEC: The Project Habitat in Greece: Netwotk Natura 2000, Life Contract B4-3200/84/756, DG XI Commission of the European Communities, The Goulandris Mouseum of Natural History, GreekBiotope/Wetland Center, Greece (in Greek) (1997)
13. Pavlikakis, G.E., Tsihrintzis, V.A.: A quantitative method for accounting human opinion, preferences and perceptions in ecosystem management. Journal of Environmental Management 68, 193–205 (2003)
14. Pavlikakis, G.E., Tsihrintzis, V.A.: Perceptions and preferences of the local population in Eastern Macedonia and Thrace National Park in Greece. Landscape and Urban Planning 77, 1–2 (2005)
15. Mara, D.D., Pearson, H.W.: Artificial freshwater environments: waste stabilization ponds. Biotechnology, VCH Verlagsgesellschaft, Weinheim 8, 177–206 (1986)
16. Mara, D.D., Pearson, H.W.:: Waste Stabilization Ponds - Design Manual for Mediterranean Europe, WHO, Regional Office for Europe, EUR/ICP/CWS 053 (1987)
17. Mara, D.D., Alabaster, G.P., Pearson, H.W., Mills, S.W.: Waste Stabilization Ponds: A Design Manual for Eastern Africa. Lagoon Technology International Ltd., England (1992)
18. Greek Official Gazette 859/86: Joint prefectural decision for the discharge of treated wastewaters in the rivers of Xanthi and Rhodope prefectures (in Greek) (1986)
19. Tsagarakis, K.P., Dialynas, G.E., Angelakis, A.N.: Water resources management in Crete (Greece) including water recycling and reuse and proposed quality criteria. Agricultural Water Management 66, 35–47 (2004)

20. WHO (World Health Organization): Health guidelines for the use of wastewater in agriculture and aquaculture. Technical Report Series 778, World Health Organization, Geneva (1989)
21. Greek Ministry of Environment, Physical Planning and Public Works: Joint ministerial Decision 5796/23-3-96. Official Gazette 854/B/16-9-1996 (in Greek) (1996)
22. Saaty, T.L.: A Scaling Method for Priorities in Hierarchical Structures. Journal of Mathematical Psychology 15, 234–281 (1977)
23. Kao, J.J., Lin, H.Y.: Multifactor Spatial Analysis for Landfill Siting. Journal of Environmental Engineering 122, 902–908 (1996)
24. Lin, H., Kao, J.: Enhanced Spatial Model for Landfill Siting Analysis. Journal of Environmental Engineering 125, 845–851 (1998)
25. Lukasheh, A.F., Droste, R.L., Warith, M.A.: Review of Expert System (ES), Geographical Information System (GIS), Decision Support System (DSS) and their application in Landfill Design and Management. Waste Management and Research 19, 177–185 (2001)
26. Sener, B., Süzen, M.L., Doyuran, V.: Landfill site selection by using geographic information systems. Environmental Geology 49, 376–388 (2006)
27. Gemitzi, A., Tsihrintzis, V.A., Voudrias, E., Petalas, C., Stravodimos, G.: Combining geographic information system, multicriteria evaluation techniques and fuzzy logic in siting MSW landfills. Environmental Geology 51, 797–811 (2007)
28. Gemitzi, A., Tsihrintzis, V.A., Christou, O., Petalas, C.: Use of GIS in siting stabilization pond facilities for domestic wastewater treatment. Journal of Environmental Management 82, 155–166 (2007)
29. Bagchi, A.: Design, Construction and Monitoring of Sanitary Landfill, pp. 235–245. Wiley, New York (1990)
30. Nemani, R.R., Keeling, C.D., Hashimoto, H., Jolly, W.M., Piper, S.C., Tucker, C.J., Myneni, R.B., Running, S.W.: Climate-Driven Increases in Global Terrestrial Net Primary Production from 1982 to 1999. Science 300, 1560–1563 (2003)
31. Gemitzi, A., Petalas, C., Tsihrintzis, V.A., Pisinaras, V.: Assessment of groundwater vulnerability to pollution: a combination of GIS, fuzzy logic and decision making techniques. Environmental Geology 49, 653–673 (2006)
32. Jenson, S.K., Domingue, J.O.: Extracting Topographic Structure from Digital Elevation Data for Geographic Information Systems Analysis. Photogrammetric Engineering and Remote Sensing 54, 1593–1600 (1998)
33. Rao, M., Sastry, S.V.C., Yadar, P.D., Kharod, K., Pathan, S.K., Dhinwa, P.S., Majumdar, P.S., Sampat Kumar, D., Patkar, V.N., Phatak, V.K.: A Weighted Index Model for Urban Suitability Assessment—A GIS Approach, Bombay Metropolitan Regional Development Authority, Bombay, India (1991)
34. Eastman, J.R., Jiang, H.: Fuzzy Measures in Multi-Criteria Evaluation. Paper presented at the 2nd International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Studies, Fort Collins, Colorado, May 21-23, pp. 527–534 (1996)
35. Yager, R.: On Ordered Weighted Averaging Aggregation Operators in Multicriteria Decision Making. IEEE Transactions on Systems, Man, and Cybernetics 8, 183–190 (1988)
36. Aivaliotis, V., Dokas, I., Hatzigiannakou, M., Panagiotakopoulos, D.: Functional Relationships of Landfill and Landraise Capacity with Design and Operation Parameters. Waste Management and Research 22, 283–290 (2004)
37. European Compost Network (2004), `http://www.compostnetwork.info` (Cited November 15, 2005)

38. Aller, L., Bennet, T., Lehr, H.J., Petty, J.R., Hackett, G.: DRASTIC: a standardized system for evaluating ground water pollution potential using hydrogeologic settings. In: Ada, O.K., Kerr, R.S. (eds.) Environmental Research Laboratory, US Environmental Protection Agency Report EPA-600/2–87–035, p. 622 (1987)
39. Foster, S.S.D.: Fundamental concepts in aquifer vulnerability, pollution risk and protection strategy. In: van Duijvenbooden, W., van Waegeningh, H.G. (eds.) Proc. Inf. of Vulnerability of soil and groundwater to pollutants, The Hague. TNO Committee on Hydrological Research, vol. 38, pp. 69–86 (1987)
40. Van Stempvoort, E.L., Wassenaar, L.: Aquifer vulnerability index: GIS-compatible method for groundwater vulnerability mapping. Canadian Water Resources 18, 25–37 (1993)
41. Civita, M.: Le carte della vulnerabilità degli acquiferiall' inquinamento. Teoria and practica (Aquifer vulnerability maps to pollution). Pitagora, Bologna (1994)
42. Civita, M., De Regibus, C.: Sperimentazione di alcune metodologie per la valutazione della vulnerabilità degli aquiferi. Q Geol Appl Pitagora, Bologna 3, 63–71 (1995)
43. Daly, D., Dassargues, A., Drew, D., Dunne, S., Goldscheider, N., Neale, S., Popescu, I.C., Zwahlen, F.: Main concepts of the European approach to karst-groundwater-vulnerability assessment and mapping. Hydrogeology Journal 10, 340–345 (2002)
44. Von Hoyer, M., Sofner, B.: Groundwater vulnerability mapping in carbonate (karst) areas of Germany, Federal institute for geosciences and natural resources, Archive no 117854, Hanover, Germany (1998)
45. Doerfliger, N., Jeannin, P.Y., Zwahlen, F.: Water vulnerability assessment in karst environments: a new method of defining protection areas using a multi-attribute approach and GIS tools (EPIK method). Environmental Geology 39, 165–176 (1999)
46. Gogu, R.C., Dassargues, A.: Current trends and future challenges in groundwater vulnerability assessment using overlay and index methods. Environmental Geology 39, 549–559 (2000)
47. Gogu, R.C., Hallet, V., Dassargues, A.: Comparison of aquifer vulnerability assessment techniques. Application to the Néblon river basin (Belgium). Environmental Geology 44, 881–892 (2003)
48. Dixon, B., Scott, H.D., Dixon, J.C., Steele, K.F.: Prediction of aquifer vulnerability to pesticides using fuzzy rule-based models at the regional scale. Physical Geography 23, 130–153 (2002)
49. Dixon, B.: Applicability of neuro-fuzzy techniques in predicting ground-water vulnerability: a GIS-based sensitivity analysis. Journal of Hydrology 309, 17–38 (2005)
50. Vergis, E.: Hydrogeological study of West Thrace and East Macedonia district. Institute of Geological and Mining Exploration, Greece (2000)
51. Papadopoulos, K., Favas, N., Anastasiadis, I., Papadopoulos, P., Gemitzi, A.: Hydrological research of Makri-Maronia area. Institute of Geological and Mining Exploration, Greece (2001)
52. Dimadis, E., Gemitzi, A., Karmis, P.: Hydrogeological study of Rodopi's district mountainous area. Institute of Geological and Mining Exploration, Greece (2001)
53. Papadopoulos, K., Romaidis, I.: Hydrogeological Research of Arda's Watershed. Institute of Geological and Mining Exploration, Greece (2002)
54. Petalas, C.P.: Analysis of aquifer systems in the heterogeneous coastal plain of Rhodope region. Doctoral Dissertation, University of Thrace, Department of Civil Engineering, Xanthi, Greece (in Greek) (1997)

55. Petalas, C., Panilas, S.C., Diamantis, J., Pliakas, F.: Calculation and evaluation of the hydraulic parameters of two aquifers in different sedimentary environment in mass of Rhodope. In: Proceedings of the 9th International Congress of the Geological Society of Greece, Athens, pp. 1913–1932 (2001)
56. Diamantis, I.: Investigation of water supply possibilities of coastal settlements of Thrace from coastal aquifers. Democritus University of Thrace, Division of geotechnical engineering, Lab. of Engineering Geology, Research project 1989 – 1993 (1993)
57. Panilas, C.S.: Hydrogeological problems from the exploitation of black coal deposits with the method of open pit: The case of the black coal mineral deposit of Drama. Doctoral Dissertation. University of Patras, Department of Geology, Patra, Greece (1998)
58. Pliakas, F., Diamantis, I., Petalas, C.: Environmental effects of a proposed groundwater artificial recharge project in Orfanon region of Kavala. In: Proceedings of the 6th International Conference on Environmental Science and Technology, Pythagorion, Samos, pp. 381–390 (1999)
59. Fetter, C.V.: Applied Hydrogeology, 3rd edn., p. 91. Macmillan College Publishing Company, New York (1994)
60. Wyssling, L.: Eine neue Formel zur Berechnung der Zustromungsdauer (Laufzeit) des Grundwassers zu einem Grundwasser-Pumpwerk. Eclogae geol. Helv. 72, 401–406 (1979)
61. Greek Ministry of Environment, Physical Planning and Public Works: Law 1650 for the Protection of the Environment. Official Gazette 160/A/1986 (in Greek) (1986)
62. National Research Council: Ground water vulnerability assessment: predicting relative contamination potential under conditions of uncertainty. Comm Techniques for Assessing Ground Water Vulnerability, Water Sci. Technol. Board, Comm. Geosci. Environ. Resources, Nat. Acad. Sci., 204 (1993)
63. Bekesi, G., Mc Conchie, J.: The use of aquifer-media characteristics to model vulnerability to contamination, Manawatu region, New Zealand. Hydrogeology Journal 10, 322–331 (2002)

Chapter 3

Integrating Multimedia GIS Technologies in a Recommendation System for Geotourism

Eleni Charou[1], Katerina Kabassi[2], Aristotelis Martinis[2], and Marianthi Stefouli[3]

[1] Institute of Informatics & Telecommunications
NCSR Demokritos, P. Grigoriou, 15310, Greece
[2] Department of Ecology and the Environment,
Technological Educational Institute of the Ionian Islands,
2 Kalvou Sq., 29100 Zakynthos, Greece
[3] Institute of Geological & Mineral Exploration
IGME building Olympic Village-Acharnai, 136 77, Greece

Abstract. In this paper a combination of traditional GIS along with multimedia, remotely sensed and GPS derived data is realised. Such a combination results in a multimedia GIS for accurate and valuable information concerning geotourism. A special feature of the multimedia GIS that was developed is that it contains information about the environmental challenges and sustainable development issues for geosites. This paper focuses on the incorporation of this multimedia GIS has been incorporated in a recommendation system for geotourism. The proposed recommendation system has the ability to personalise its interaction to each individual user. In this respect the user is proposed to visit the touristic sights that have geological interest and seems to be of his/her interests, needs and knowledge.

1 Introduction

Geotourism is based on the magic of discovery and the power of authenticity experienced through the contact with the natural heritage. Geotourism is a form of cultural-environmental tourism that can be developed in areas with important geological monuments, which are exploited in order to attract visitors with special interests. These monuments are called Geotopes or geosites. The Geotopes are the meeting places of elements recording the geological history of each region, witnessing the everlasting evolution of life on Earth. Sites such as caves, gorges, fossilized areas, large geological rifts, ancient mines, geological formations or landscapes chiselled by natural forces throughout the geological ages have a special scientific and aesthetic value and could become areas with significant tourist interest.

The Greek region constitutes one of the most significant environmental and cultural reserves on Earth, strewn with unique and significant natural geological monuments. These geotopes have different characteristics and attract different kind of tourists.

G.A. Tsihrintzis and L.C. Jain (Eds.): Multimedia Services in Intel. Environ., SIST 3, pp. 63–74.
springerlink.com

However, finding the geotope that would fit best a user's interests, knowledge and needs is a rather difficult task. This task is mainly addressed by recommender system. Indeed, recommender systems for tourism have attracted a lot of interest lately as the touristic industry expands rapidly.

In view of the above we have run the first phases of software's life-cycle for the development of a recommendation system for geotourism in Greece. A software life-cycle consists of four phases: Requirements specification, Design, Implementation and Testing. In this paper we mainly focus on Requirements specification, Design and some implementation issues of the recommendation system for geotourism in Greece.

The main characteristic of the proposed recommender system is that it provides personalized recommendations. Lately personalised recommendation systems have been gaining interest in tourism to assist users with their travel plans (Ricci 2002, Ricci & Werthner 2002, Wallace et al. 2003, Loh et al. 2003). Personalisation involves the design of enabling systems to capture of infer the needs of each person and then to satisfy those needs in a known context (Riecken 2000). A personalisation system is based on three main functionalities: content selection, user model adaptation and presentation of results (Mizarro & Tasso 2002, Diaz & Gervas 2005, Diaz et al. 2008). By content selection, one may refer to selecting destination, tourist attractions, accommodations, routes or all the above for planning a whole trip. Most of the recommendation systems, like the one described in this paper, focus on selecting the destination except from a few exceptions (Ardissono et al. 2003, Niaraki & Kim 2009).

In order to choose the destination that would fit best the user's interests and requirements, the recommendation system should incorporate a user model. Several approaches have been proposed in the international literature for the elaboration of this information in order to evaluate the alternative destinations. For example, Huang and Bian (2009) uses AHP theory (Saaty 1980) and Chin & Porage (2001) the Multi Attribute Utility Theory (Vinke 1992) for evaluating the available travelling opportunities and proposes the user the one that fits best his/her needs and preferences. In our approach we use a simple multi-criteria decision making theory, Simple Additive Weighting (SAW) (Fishburn 1967; Hwang and Yoon, 1981) in combination with stereotypes and individual user modelling for modelling the users' preferences.

Another characteristic of some recommender systems is that they incorporate GIS to improve the presentation of the recommendations and, therefore, increase the interactivity of their systems. The project CReation of User-friendly Mobile services Personalised for Tourism (CRUMPET) (Posland et al. 2001, Laukkanen et al. 2002, Schmidt-Belz et al. 2002; Zipf 2002) and the SPETA system (Garcia-Crespo et al. 2009) use GIS in order to provide user-friendly mobile services for tourism. A quite different approach is employed by Niaraki & Kim (2009), who use a GIS in a personalised ontology-based route planning system. However, the approach described is quite different as it uses a multimedia GIS and the domain of geotourism is a special case of tourism.

The use of GIS as a basis for a system related to geotourism is reasonable as the majority of touristic information is spatially related. The GIS system allows for the combined presentation of various data types and allows the user to determine the distances between various points of interest. On the other hand, a very popular and widely used digital form of data is multimedia. Multimedia data types i.e. text, images, audio and video and provides such an atmosphere in which user feels comfortable with the system environment before addressing issues of data analysis. The combination of multimedia along with traditional GIS can improve the performance of GIS turning it into a more realistic tool for spatial analysis that has full capability to provide services to both technical and non-technical users. This will allow more clear and transparent view of overall situation and better understanding of its spatiotemporal implications. These characteristics initiated various Mutlimedia GIS (MMGIS) implementations (Bakourou et al. 2002), (Ayeni et al. 2004), (Bordoni & Colagrossi 1999), (Frech et al. 2003), (Papadimitriou & Roustanis, 2007). The integration of Multimedia and GIS can be viewed in two ways (Soomro et al. 1999): a) “GIS in Multimedia”: where multimedia-authoring tools are used to blend different data types (text, image, video, audio) and spatial analytical function to boost traditional GIS or in other words MM has capabilities to incorporate GIS and b) “Multimedia in GIS” where the traditional GIS will encompass the capabilities to take care of multimedia or in other words GIS has capabilities to incorporate MM data types. The MMGIS developed in this study falls to the second category. The TNTmips V 7.3 that supports fully integrated GIS, image processing, CAD, TIN, Desktop cartography and geospatial database management has been used for the processing of data and the MMGIS implementation.

2 Requirements Analysis and MMGIS Database Design

Tourism in general is not just business but has also to do with emotional experiences; [Werthner, 2001] and lifestyle. Geotourism, in particular, can also improve a visitor’s knowledge and environmental consciousness. In order to capture requirements for the recommender system that incorporates the MMGIS, an empirical study was conducted. The empirical study participated tourists of different interests, needs and knowledge that were visiting two different Geotopes, the island of Zakynthos and the wider region of Micro and Macro Prespa lakes in Greece (Fig. 1). Both areas are of great ecological importance and constitute geotopos.

Zakynthos has attracted the international interest because the loggerhead turtle *Caretta-caretta*, one of the endangered species protected by international conventions and by Greek legislation, lays its eggs on the island. Zakynthos has a great number of geosites representing active and evolving geomorphological landforms, including tectonically active fault scarps, geothermal fields, karst and caves, and coastal and fluvial landforms. The Prespa region, on the other hand, is also considered to be an ecosystem of global significance and has been identified as one of Europe's 24 major transboundary "ecological bricks". The area is quite mountainous with interesting geomorphological features, important mines and lakes.

During the empirical study, information characterizing current and future needs of geotourism community was gathered. The study also revealed that people interested in Geotourism or other kinds of natural tourism are also interested in current environmental challenges and sustainable development. Therefore, the system should address these new trends along with providing potential users with valuable, accurate and up to date information regarding the geosites. In this respect the GIS would prove to be very useful. More specifically, the system should be designed to address various environmental issues such as biodiversity, land cover changes, deforestation, water pollution, etc.

Remotely sensed data constitute a valuable tool for obtaining such environmental information. A special interest was given to the way of data visualization using 3D views, virtual flights and panoramas. 3D renderings could also have considerable impact in the way the geosities are perceived.

Fig. 1. Overview of the study area.

The data that were gathered, were categorized: Each category constitutes a layer or a sub layer in the designed GIS database:

Geological: Geological / geophysical / geomorphological features, mountains, lineaments, cliffs, waters, falls, springs, beaches.
Ecological: beaches, national parks, protected areas, forest reserves, botanical gardens etc.
Cultural: Museum, art collections, cultural, religious and national monuments.
Other: sporting facilities, travel and accommodation facilities, tourism Centers and agents, and other related to tourism data.

Sources of data include existing maps, plan and charts, textbooks, and Internet websites. The existing map data have been converted into digital form. The old map data were used in order to compare the possible changes that have taken place in the area. Satellite images were used for visualization due to their ability to show spatial patterns of land cover. Combinations of different resolution data using data fusion techniques based on HIS colour model proved to be effective as far as the interpretation of features of interest is concerned (Charou et al. 2003). Satellite images were processed and new image products addressing the detection of changes of land cover were generated.

3D views constitute a novel form of information presentation. They are very suitable for the visualization of a region as they give the user a particularly realistic impression. They provide information not only about distances but also about gradients along a chosen route. 3D views as well as virtual flights and panoramas permit users to view the area from a variety of angles giving them a comprehensive overview of the region (Almer et al. 2000). In order to generate 3D views, virtual flights and panoramas as realistic as possible, Digital Elevation Models (DEMs) were prepared and integrated with satellite data.

Walking and cycling are amongst the most popular activities associated with geotourism. Accurate information concerning geotouristic bike and/or hiking routes are of a great importance for these activities. To this end, GPS surveys on selected routes were carried out and GPS data were integrated with satellite data and Digital Elevation Models for a better representation.

Digital cameras and digital video cameras were also used to capture data. Video clips with sound were hot-linked with the other types of data in the TNT mips environment using Script files thereby creating a multimedia GIS database containing digital photos, audio and video files.

3 Design of the Recommendation System

Travel preferences are often hidden and are not explicitly known when users start to plan their trips, particularly if visiting an unfamiliar place (Viappianni et al. 2002; Loh et al 2003). The system should contain a large amount of detailed up-to-date information about the possible destinations (Fesenmaier & Jeng 2000, Ardissono et al. 2003). Therefore, the system contains information about the places of Greece that have geological interest as well as related information about each place of interest.

The system architecture consists of:

- The User Modeling component, which initializes, updates and maintains information about each user and the stereotypes
- The Recommender, which selects the place of interests and ranks them to suggest the best one. The selections process in the recommendation system consists of two stages: estimating travelers' preferences, and subsequently evaluating available destinations.
- The Personalisation component, which is responsible for dynamically updating the pages of the web store.
- The GIS module that is responsible for presenting the information in the GIS.

In order to estimate the traveler's preferences, the system incorporates a user modelling component. This component maintains information about the interests, needs and background knowledge of all potential users. As Kobsa (2001) point out, a user modeling system has to carry out the following tasks:

- Learn the interests and preferences of users based on their usage of the application
- Predict interest and preferences of individual users based on assumptions about homogeneous user subgroups (stereotypes)
- Infer additional interest and preferences using domain knowledge
- Store, update and delete explicitly provided information and implicitly acquired assumptions.

Taking into account these requirements, the user modelling component of the proposed system maintains information about each individual user that interacts with the system. Additionally, the system uses stereotypical knowledge for modelling the users for whom it has not gathered enough information yet. As Rich (1989, 1999) points out a stereotype represents information that enables the system to make a large number of plausible inferences on the basis of a substantially smaller number of observations. Stereotypes constitute a common user modelling technique for drawing assumptions about users belonging to different groups and, therefore, it has been used in a wide range of applications such as e-learning systems (Alepis & Virvou 2006, Kabassi et al. 2006), help systems (Virvou & Kabassi 2004), recommendation systems (Virvou et al. 2006, Chin & Porage 2001) etc.

More specifically, the user modeling component of the system maintains information about a user's interests, economical preferences, time constraints and background knowledge. The first three criteria have been referred in many other recommender systems for tourism. However, background knowledge is a new criterion that has been added in this system due to geotourism. The user should have the knowledge to understand the geological phenomenon that exists in the evaluated and selected destination. The information that is stored in the user model is acquired explicitly by the user as it is beyond the system's scope to check the users' knowledge on the subject of geology or related research areas. However, information that involves the interests of the particular user is acquired explicitly in the beginning and this information is constantly updated during the user's interaction with the system.

The system uses a content-based filtering for personalization. Content-based filtering suggests to a user items or services which are similar to those he/she bought or searched in the past, by matching the characteristic of the item or service with the characteristics of the user that is maintained in his/her user model. The main advantage of content-based filtering is that it is based only on facts that involve the particular user and, therefore, are true. Of course, in some cases this advantage may be also disadvantage as the particular method may lead to overspecialized suggestions that include only items excessively similar to those the user already knows (Adomavicius and Tuzhilin, 2005). Another advantage of content-based filtering that derives from the fact that it is based on information for each individual user is that it may also capture changes on the user's preferences.

Another aspect of the recommendation system involves the enhancement of it decision making process. Ricci (2002) stated that a recommender system could provide valuable information to assist consumers' decision-making process. The decision making theories that seem to be more appropriate for computer problems are the multi-criteria ones. This is due to the fact that computer problems usually involve several objectives and criteria. Decision making theories provide precise mathematical methods for combining criteria and rank the alternatives from best to worst and determine a subset of actions considered to be the best in order to make the final decision.

As a result different decision making theories have been used in recommendation systems for tourism. For example, Chin & Porage (2001) the Multi-Attribute Utility theory (MAUT) (Vinke 1992) in combination with stereotypes to evaluate the available travelling opportunities and proposes the user the one that fits best his/her needs and preferences. The system described in this paper uses Simple Additive Weighting (SAW) (Fishburn 1967, Huang & Yoon 1984) for evaluating the different alternatives and select the one that seems to be the most appropriate for the particular user.

4 Implementation Issues of Multimedia GIS

A number of geological features were digitized from the Geological maps of the Institute of Geological & Mineral Exploration IGME (scale 1: 50.000). More specifically, geologic layers (vector) containing the hydrological network, lithological unit boundaries and tectonics (faulting and bedding system) were created.

A similar procedure was followed in the digitization of the soil maps (land use and land capability for forestry) of the Ministry of Agriculture (scale 1: 20.000) and Topographic Maps of the Geographic Service of the Army (scale 1:50.000). Topographical data include the coastline, the main and secondary road network, caves, meteorological stations and village polygons (outline of village limits). Following the digitization of the maps, georeferencing of them was performed with TNT mips software, by choosing specific GCPs in the corresponding maps and the digitized coastline.

Additionally, multisensor satellite data such as LANDSAT-ETM, SPOT, ASTER and MERIS data having various acquisition dates have been processed. In order to prepare these satellite images for further processing, geometric and radiometric corrections were performed. When necessary, data have been re-projected into the local Greek Geodetic system georeference -Egsa 87. For MERIS data georeferencing has been made using the tie-points provided with the images. In order to combine different resolution data, data fusion techniques were used. The multitemporal satellite data were analysed in order to indentify and present changes of land cover during the last 15 years.

Digital elevation models for the areas of interest based on fused LANDSAT and SPOT with 10m spatial resolution RGB satellite data were generated and used for the construction of 3D views (Fig. 2). The ideal way to visualize geospatial data in 3D is to "fly" over and around a 3D scene interactively and examine features from any direction in real time. Virtual flights were generated by selecting an elevation raster (or other surface raster) to define the terrain and optionally selectthe same raster or an image of the same area to use as an overlay (drape layer) (Fig. 3).

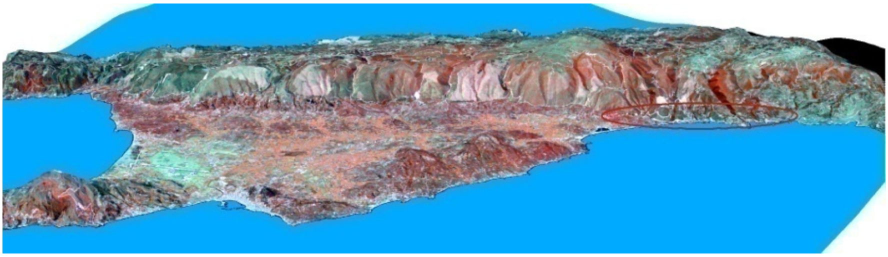


Fig. 2. A 3D View of the Zakynthos island.

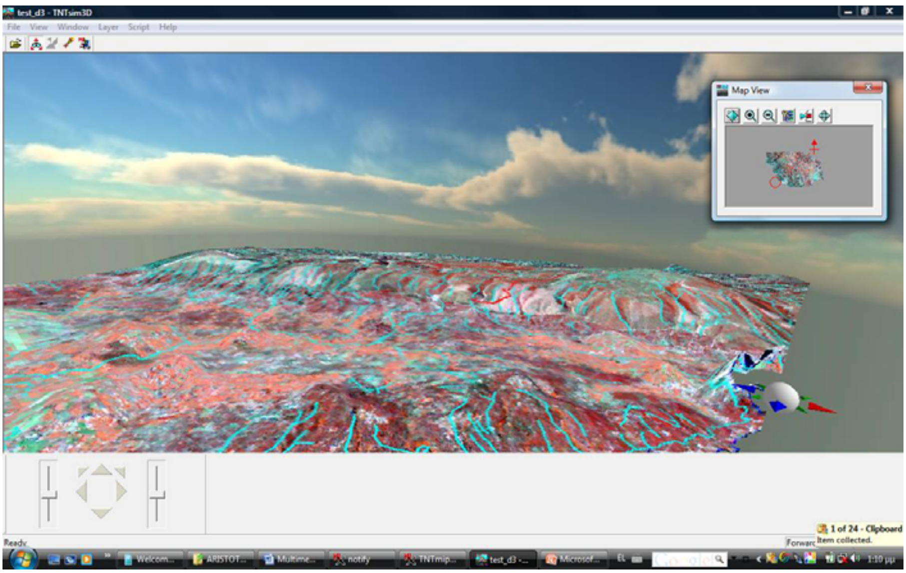


Fig. 3. A Screenshot of a virtual flight over the Zakynthos island.

A Thales Navigation Mobile Mapper GPS unit with post processing capabilities was used to delineate various footpaths, routes and points of interest. Mapping was performed when ideal satellite and PDOP numbers were available. The goal of successful mapping was to have the highest number of satellites combined with the lowest number for the PDOP. The minimum number of satellites that could enable accurate mapping was five. The highest number PDOP allowable for mapping was eight. At the end of each survey, data was uploaded into the computer and exported into the TNT mips. More specifically, for every footpath, several maps and various diagrams related to the description of the land cover, the landscape, the degree of difficulty (slopes) and distances were created.

The available satellite data offers the possibility to identify the surface extent of the lakes of the region and to map the coastlines of various lakes in different dates (Charou et all 2004). On the basis of the available maps and satellite data, digital elevation models were used in order to delineate various features (Fig. 4). Additionally, MERIS data have been proven capable of estimate land cover changes (Stefouli et al. 2008, Charou et al. 2004) in the broader area of Prespa region.

Fig. 4. Selected GIS maps of Prespa region

The analysis also revealed the devastating effect that forest fires had in Western Balkan peninsula during summer 2007 (Fig. 5). Significant land cover changes occurred in the area due to fire events and these were detected using MERIS data. A superposition of Corine 2000 reveals that significant land cover changes have occurred between the years 2000 to 2007.

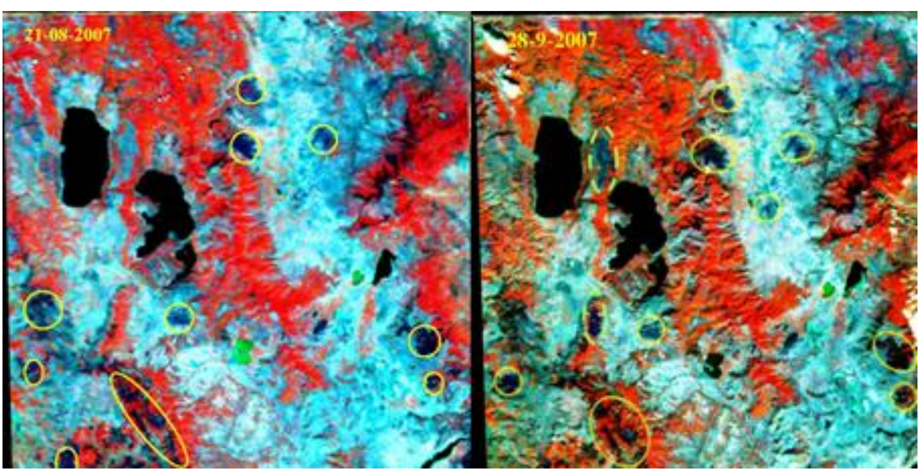


Fig. 5. Burnt areas (yellow curves) in Prespes test area using MERIS data.

Finally field –collected Multimedia data containing digital photos, audio and video files were linked to their locations and used to populate the mmgis database. Video was used to show background, point scenes, or transition and audio to provide realism by communicating the notion of space. The link between geographic features and multimedia ïïïïï is done using script files. All data have been organized in a form of an informational Atlas.

5 Conclusions

In this paper, we describe the way that new and innovative technologies such as 3D views and multimedia were combined with traditional GIS to develop a MMGIS for geotourism. The combination of data from different sources linked by the powerful prospects of new Earth Observation data and techniques resulted in the development of new knowledge that can be used for ïï ï ï ï ïïïï ï geotourism. For this purpose, it has been used in a recommendation system for geotourism.

For the system's requirements specification, an empirical study with the participation of tourists in geosites was conducted. The empirical study revealed the kind of information that would attract the possible end users of a recommender system for geotourism as well as the individual characteristics of such users. Therefore, during the design phase of the system the architecture and the personalisation procedure of the system were specified. In order to estimate the traveler's preferences, the system incorporates a user modeling component. This component maintains information about the interests, needs and background knowledge of all potential users. This information is further used by the system in order to select the kind of information that will present to the user as well as the geosites that is possible to interest him/her. Then, it provides the geotourism community accurate and valuable information for the geosites. For this purpose, up to date satellite images, maps and multimedia data were used. The system can be used both in computers and handheld devices.

The proposed system could also be used as a tool to learn about current environmental challenges and it is designed to address sustainable development issues including biodiversity, land cover changes, deforestation, water pollution, etc. The proposed approach could easily be modified to reach various target groups for other forms of natural tourism (eco, adventure) or Cultural Tourism (rural, popular cultural heritage - events, religious and product) tourism.

Acknowledgment

European Space Agency is acknowledged for providing the MERIS data in the framework of ESA Cat.-1 project no. 4864.

References

1. Alepis, E., Virvou, M.: Emotional intelligence: Constructing user stereotypes for affective bi-modal interaction. In: Gabrys, B., Howlett, R.J., Jain, L.C. (eds.) KES 2006. LNCS (LNAI), vol. 4251, pp. 435–442. Springer, Heidelberg (2006)

2. Almer, A., Nischelwitzer, A.K.: 3-D Visualisation of Leisure & Tourism Information Based on Remote Sensing Data; Proceedings of ISPRS Congress, Amsterdam, The Netherlands, CD-ROM, TC V-05-03, 1,337 (2000)
3. Ayeni, O.O., Saka, D.N., Ikwuemesi, G.: Developing a Multimedia Gis Database for Tourism Industry in Nigeria ISPRS Insanbul, Turkey, pp. 630–635 (2004)
4. Ardissono, L., Goy, A., Petrone, G., Signan, M., Torasso, P.: Intrigue: Personalized recommendation of tourism attractions for desktop and handset devices. Applied Artificial Intelligence 17(8-9), 687–714 (2003)
5. Bakourou, E., Tsioukas, V., Katzougraki, I., Stylianidis, E., Papadimitriou, K., Patias, P.: The promotion of cultural heritage through internet using advanced audio-visual information: The venetian castles of Peloponnisos. In: IAPRS&SIS, Com V, Corfu, Greece, vol. XXXIV, Part 5, pp. 298–301 (2002)
6. Bordoni, L., Colagrossi, A.: A multimedia personalized fruition of figurative artistic heritage by a GIS-based methodology. In: IEEE International Conference on Multimedia Computing and Systems, vol. 2, pp. 184–188 (1999)
7. Charou, E., Petridis, S., Stefouli, M., Mavrantza, O.D., Perantonis, S.J.: Innovative feature selection used in multispectral imagery, classification for water quality monitoring. In: Proceedings of ISPRS Congress, Istanbul, Turkey (2004)
8. Charou, E., Vassilas, N., Perantonis, S., Varoufakis, S.: Intelligent Techniques in Environmental Data Processing. In: Harmancioglu, et al. (eds.) Integrated Technologies for Environmental Monitoring and Information Production, pp. 301–310. Kluwer Academic Publishers, Dordrecht (2003)
9. Chin, D., Porage, A.: Acquiring User Preferences for Product Customization. In: Bauer, M., Gmytrasiewicz, P., Vassileva, J. (eds.) UM 2001. LNCS (LNAI), vol. 2109, pp. 95–104. Springer, Heidelberg (2001)
10. Diaz, A., Gervas, P.: Personalisation in news delivery systems: Item summarization and multi-tier item selection using relevance feedback. Web Intelligence and Agent Systems 3(3), 135–154 (2005)
11. Diaz, A., Garcia, A., Gervas, P.: User-centred versus system-centred evaluation of a personalisation system. Information Processing and Management 44, 1293–1307 (2008)
12. Fesenmaier, D.R., Jeng, J.: Assessing structure in the pleasure trip planning process. Tourism Analysis 5(3), 13–28 (2000)
13. Fishburn, P.C.: Additive Utilities with Incomplete Product Set: Applications to Priorities and Assignments. Operations Research 15, 537–542 (1967)
14. Frech, I., Koch, B.: Multimedia Geoinformation in Rural Areas with Eco-tourism: The ReGeo-System. In: Frew, A.J., Hitz, M., O'Connor, P. (eds.) Proceedings of the International Conference in Information and Communication Technologies in Tourism, Helsinki, Finland, pp. 421–429. Springer, NewYork (2003)
15. Garcia Cerspo, A., Chamizo, J., Rivera, I., Mncke, M., Colomo Palacios, R., Gomez Berbis, J.M. SPETA: Social pervasive e-Tourism advisor. Telematics and Informatics (2009) (to appear)
16. Huang, Y., Bian, L.: A Bayesian network and analytic hierarchy process based personalized recommendations for tourist attractions over the Internet. Expert Systems with Applications 36(1), 933–943 (2009)
17. Hull, M.E.C., Taylor, P.S., Hanna, J.R.P., Millar, R.J.: Software development processes-an assessment. Information and Software Technology 44, 1–12 (2002)
18. Hwang, C.L., Yoon, K.: Multiple Attribute Decision Making: Methods and Applications. Lecture Notes in Economics and Mathematical Systems, vol. 186. Springer, Heidelberg (1981)

19. Kabassi, K., Virvou, M., Tsihrintzis, G.A.: Hybrid Intelligent Medical Tutor for Atheromatosis. In: Gabrys, B., Howlett, R.J., Jain, L.C. (eds.) KES 2006. LNCS (LNAI), vol. 4252, p. 1289. Springer, Heidelberg (2006)
20. Laukkanen, M., Helin, H., Laamanen, H.: Tourists on the move. In: Klusch, M., Ossowski, S., Shehory, O. (eds.) CIA 2002. LNCS (LNAI), vol. 2446, pp. 36–50. Springer, Heidelberg (2002)
21. Loh, S., Lorenzi, F., Saldana, R., Licthnow, D.: A tourism recommendation system based on collaboration and text analysis. Information Technology & Tourism 5, 157–165 (2003)
22. Mizzaro, S., Tasso, C.: Ephemeral and persistent personalization in adaptive information access to scholarly publications on the web. In: De Bra, P., Brusilovsky, P., Conejo, R. (eds.) AH 2002. LNCS, vol. 2347, pp. 302–316. Springer, Heidelberg (2002)
23. Niaraki, A.S., Kim, K.: Ontology based personalized route planning system using a multi-criteria decision making approach. Expert Systems with Applications 36, 2250–2259 (2009)
24. Posland, S., Laamanen, H., Malaka, R., Nick, A., Buckle, P., Zipf, A.: Creation of user-friendly mobile tourism services personalised for tourism. In: Proceedings of the 2nd International Conference on 3G Mobile Communication Technologies, London, UK, pp. 28–32 (2001)
25. Ricci, F.: Travel recommendation systems. IEEE Intelligent Systems 17(6), 55–57 (2002)
26. Ricci, F., Werthner, H.: Case-based querying for travel planning recommendation. Information Technology and Tourism 4(3-4), 215–226 (2002)
27. Rich, E.: Stereotypes and User Modeling. In: Kobsa, A., Wahlster, W. (eds.) User Models in Dialog Systems, pp. 199–214 (1989)
28. Rich, E.: Users are individuals: individualizing user models. International Journal of Human-Computer Studies 51, 323–338 (1999)
29. Riecken, D.: Personalised views of personalization. Communications of the ACM 43(8), 27–28 (2000)
30. Saaty, T.L.: The analytic hierarchy process. McGraw-Hill, New York (1980)
31. Schmidt-Belz, B., Polsad, S., Nick, A., Zipf, A.: Personaliszed and location-based mobile tourism services. In: Workshop on Mobile Tourism Support Systems, pp. 18–20 (2002)
32. Virvou, M., Kabassi, K.: Adapting the Human Plausible Reasoning Theory to a Graphical User Interface. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans 34(4), 546–563 (2004)
33. Soomro, T.R., Zheng, K., Pan, Y.: Capabilities of multimedia gis Chinese. Gegraphical Science 9(2), 159–165 (1999)
34. Stefouli, M., Kouraev, A., Charou, E.: Monitoring water level and water quality of lakes Macro Prespa and Ochris using radar altimetry and MERIS data. In: 2nd MERIS A(A)TSR workshop, Frascati, Italy (2008)
35. Virvou, M., Savvopoulos, A., Tsihrintzis, G.A., Sotiropoulos, D.N.: Constructing Stereotypes for an Adaptive e-Shop using AIN-based Clustering. In: 8th International Conference on Adaptive and Natural Computing Systems, Warsaw, Poland, pp. 11–14 (2007)
36. Wallace, M., Maglogiannis, I., Karpouzis, K., Kormentzas, G., Kollias, S.: Intelligent one-stop-shop travel recommendations using an adaptive neural network and clustering of history. Information Technology & Tourism 6, 181–193 (2003)
37. Viappiani, V., Pu, P., Faltings, B.: Acquiring user preferences for personal agents. In: Proceedings of the AAAI fall symposium, pp. 53–59 (2002)
38. Vincke, P.: Multicriteria Decision-Aid. Wiley, Chichester (1992)
39. Zipf, A.: User-adaptive maps for location-based services (LBS) for tourism. In: Proceedings of the 9th International Conference for Information and Communication Technologies in Tourism, Innsbruck, Austria, vol. 1, pp. 183–197 (2002)

Chapter 4

Lighting Design Tools for Interactive Entertainment

Magy Seif El-Nasr[1] and Joseph Zupko[2]

[1] Simon Fraser University
[2] Pennsylvania State University

Abstract. Interactive entertainment is one of the fastest growing industries in the world, collecting more than $1 billion in revenues. According to the first quarter Video Gaming report published in 2005, gamers have expanded beyond the expected 8-34 old male demographic to include women, Hispanics, and African Americans. Interactive entertainment is also flourishing in global markets, including North America, Japan, China, and Korea. Recently, India has also entered this market with their first international release of a AAA 3D game title *Ghajini –The Game*. For such a global and fast growing industry, the development of better tools and processes to facilitate the development cycle is becoming an important problem. A typical AAA video game production cycle is around 6 years of development with development budgets of $50+ million dollars. While research in developing production tools has been underway for areas such as animation, writing, character scripting, and event scripting, other areas such as lighting design have received very little attention. This chapter discusses our research exploring the design and development of production tools for video game lighting, starting with our first lighting design tool the *Expressive Lighting Engine (ELE)* and finishing with our latest lighting design tool the *System for Automated Interactive Lighting (SAIL)*. We will describe these tools in detail and present a research study discussing how such tools were received by visual designers. The chapter ends with speculations on the future and the utility of such tools.

1 Introduction

Interactive entertainment is one of the fastest growing industries in the world. In 1996, the U.S. Entertainment Software Association (ESA) reported $2.6 billion in sales revenue; this figure more than tripled in 2007, yielding $9.5 billion in revenues [1]. In addition, the target market for interactive entertainment products now reaches beyond the traditional 8-34 old male demographic to include women, Hispanics, and African Americans [1, 2]. This pattern has been observed in several markets, including Japan, China, Korea, and India (who has just published its first international AAA title, *Ghajini – The Game* [3]).

With such a wide appeal for interactive entertainment products, the industry is facing several challenges, including game design innovation, management of development

G.A. Tsihrintzis and L.C. Jain (Eds.): Multimedia Services in Intel. Environ., SIST 3, pp. 75–102.
springerlink.com

teams that reach 200+ developers, and tool development. A typical game production cycle for a AAA title is around 6 years of development with development budgets of $50+ million. While research in developing production tools has been underway for areas such as animation, writing, character scripting, and event scripting, lighting design has received very little attention. This chapter concentrates on this problem. In particular, it discusses innovative and novel lighting design tools.

Lighting is an important element of any media production as it shapes all visual perception. Film, theatre, and many other design disciplines devote much energy to controlling light. In some cases the motivation is simple visibility, but lighting design is more than just maintaining visibility. Lighting can control mood, evoke emotion, and guide visual interest and intent. Knowing these important functions of light, cinematic and theatric lighting designers have developed techniques to continuously modulate scene lighting to achieve desired effects [4].

Lighting is as important for video games as it is for the film, theatre, and animation. Game lighting techniques have adopted many of the practices of theatre, film, and architecture [5, 6] to achieve similar levels of quality. However, lighting a game is different than lighting a film or a theatre production due to the unpredictability present in game environments induced by user interaction. In film, theatre and animation, the designer has omniscient knowledge of the script, the set, camera positions, and character positions and actions. This knowledge is used to carefully plan and implement every moment of a cinematic or theatric work. In games, however, this knowledge is not available. A designer does not know all the different character positions, camera positions, or narrative choices a user might make at runtime. This necessitates the development of a new lighting design process for games.

Although film and theatre lighting design techniques have been documented in many books available to non-experts, the lighting design process for games has not received such attention. In order to reveal the process that is used by game designers within the industry, we interviewed several lighting design professionals as well as art directors and level designers. The full study is discussed in [7]; we will discuss some parts of this study here to support this chapter's arguments. In total, we interviewed 19 lighting experts, 11 of which have worked on AAA titles within the game industry (see section 8 for a summary of the backgrounds of participants). In our one-on-one interviews, we asked them to specifically talk about the lighting design process and the role of lighting in the productions they have worked on.

All interviewees identified that lighting design is an important but complex process that plays different roles depending on the game genre. These roles include directing visual attention, providing mood, enabling visibility, and providing a realistic look for the level or game. For example, one interviewee mentioned that lighting is important to direct a player's visual attention to help solve spatial puzzles [7].

Current methods for lighting design in games differentiate between lighting the environment and lighting the characters within the environment. To light the environment, lighting designers usually pre-bake the lighting into textures using algorithms that attempt to create a realistic look for the environment, such as *radiosity* [6]. Characters and objects in contrast are lit dynamically in order to accommodate user interaction. Lights may be attached to objects so that they follow them around to maintain visibility or lights may be added and removed by events that designers determine in advance should cause a lighting change. This process has two fundamental problems:

1. It is time consuming as it requires designers to balance a large number of lights used for the environment against a separate set of lights used for characters, maintaining visual consistency and aesthetics between the two.
2. This lighting design model itself is still not adaptable to an unpredictable game environment. Designers must manually configure and script lighting to try to accommodate any user interaction that they can anticipate. For example, even if the obvious entrance to a building is the front door, a designer must anticipate that a player will enter through a window and design the lighting to accommodate this possibility as well.

A solution to these issues is to develop a lighting design system that can automatically adjust to user interaction at runtime, alleviating a designer from manually accommodating every anticipated player choice in advance. Such a system would produce more robust lighting designs that take less time for a designer to create.

In the past few years we have developed several such systems. In particular, we first developed the *Expressive Lighting Engine (ELE)* [8], published in 2003. ELE uses a constraint optimization[1] algorithm developed based on film and theatre lighting techniques to accommodate context while adapting to a designer's desired effect. While ELE received much success within the research and academic communities, designers who have interacted with it responded with confusion as they could not adequately understand the interdependencies between the numerical parameters they were given to set their desired lighting designs.

As a result of their evaluations and comments, we developed a new lighting design tool, the *System for Automated Interactive Lighting (SAIL).* The goal of SAIL is to adapt to changes in context given a designer's aesthetic and functional intent (specified by an image of the desired lighting). SAIL tries to achieve the best possible compromise between what the lighting should naturally look like given context and what the lighting should ideally look like as represented by an image. This system addresses ELE's shortcomings in two ways. First, it uses image analysis to extract lighting cues from an image. This removes the need for designers to understand and configure confusing numerical terms and instead allows them to use an image to encode their design constraints. Second, ELE did not understand how an object responds to lighting. To ELE, the trunk of a tree was equivalent to a human of the same size and girth. This required designers to manually tweak ELE's output to accommodate different objects. SAIL uses an image analysis algorithm in a preprocess step to learn how a 3D object responds to light. It then uses this model to adapt lighting based on the object's appearance.

As previously mentioned, video game lighting design is not as openly documented as film or theatre. However, the game industry already has set processes and pipelines. This is a highly overlooked problem that complicates the exploration and development of new lighting tools and techniques. To alleviate this problem and to evaluate the contributions of SAIL and ELE, we conducted a research study with lighting experts. We present some early results from this study later in the chapter.

The rest of this chapter is organized as follows. First, we describe the lighting design process in games and traditional media. Second, we discuss related work.

[1] Constraint optimization is a process of problem solving by maximizing or minimizing a specific mathematical function through manipulation of some parameters.

Third, we discuss ELE in more depth to set the context for our latest tool SAIL. We then discuss the algorithms of SAIL. As mentioned above, the acceptance of such a system within the industry is still a big issue that we want to explore. In an attempt to start this exploration, we will present results from a qualitative study we conducted to evaluate SAIL's contributions and the perception of such tools by professional lighting designers. We conclude by discussing speculation of the state of the art of such tools within the interactive entertainment industry and their utility.

2 Traditional Lighting

Adaptive lighting design is primarily targeted at interactive contexts such as video games. However, the fundamentals of game lighting come from more traditional domains such as film and theatre. In these domains, the role of a lighting designer is to balance several goals, such as visibility, depth, modeling, and to support the story and evoke mood and emotion [9].

In creating a lighting design, a designer considers the environment, the characters (or other objects of interest), and the narrative. She also considers time of day, the style of the work, and the mood or theme [5, 10]. A designer must balance her design goals with an understanding and respect of the physical nature of light. Understanding the real world behavior of light attenuation, light reflection, occlusion, and other phenomena allows a designer to understand when and how she can "cheat" these phenomena to accomplish her visual goals.

One common technique used in both theatre and film is known as *three-point lighting*. This method divides character lighting into three lights or sets of light: a *key* light, a *fill* light, and a *back* light [5, 10, 11]. The key light provides dominant direction cues, the fill reduces visual contrast, and the back light isolates the object, punctuating its silhouette. This lighting method allows a designer to deliberately control lighting while still reflecting physical reality. The fill light, for example, takes the place of bounce illumination through careful control of lighting instruments and exposure, allowing a designer to reflect the physical property of indirect illumination while maintaining desired control of an object's appearance.

3 Game Lighting

Game lighting process is not as well documented as film or theatre process. The discussion in this section is based on interviews we conducted with 19 lighting experts. 11 of these experts were game developers who have worked on AAA titles and causal games. They include lighting designers, art directors, and level designers.

Generally, the process of lighting design for a game starts with a visual concept, an image or collection of images that represent the desired aesthetics and overall look for a game. Lighting designers then work to replicate this concept within the 3D environment developed for the game. In this step, lighting designers often use tools such as *Autodesk 3ds Max* or *Autodesk Maya*, where they position each light manually and set its properties (e.g., angle, color, attenuation, and type). This lighting configuration is then typically baked into a game environment or level and remains static at runtime. This process usually involves several iterations using guidance and feedback from the art director. Once a lighting design is established for a given environment, designers move to the next environment. This process continues until all environments for the game are lit.

Lighting characters and other dynamic objects differs with the company, game, and technology available. Lights may be attached to an object so that they follow the object and achieve a specific look independent of context. These lights are then blended with an approximation of the static lighting used on the environment to keep the look of the object consistent with its environment. Occasionally, the same lighting technology is used for both the static environment and dynamic objects. In this case, designers will tweak the environment lighting so that it is also acceptable as object lighting.

The lighting process can also differ depending on the type of game, e.g. story-based or puzzle-based. For example, in games such as *Prince of Persia* or *Tomb Raider*, lighting is important to cue the player on clues for solving spatial puzzles. Aesthetic lighting initially designed to tell a story might later be overridden by lighting necessary to help a player navigate the world during the design process.

4 Related Work

While there is no prior work that targeted automated lighting design tools for video games, there has been similar research for other domains. This body of work is known as *inverse lighting*. Here we present a brief overview of this field. For a more complete survey, see [12].

We divide inverse lighting into three sub fields. *Architectural* inverse lighting targets visualization of real world spaces. *Perceptual* inverse lighting targets the automatic illumination of objects to a perceptual ideal. *Interactive* inverse lighting targets interactive, real-time lighting.

4.1 Architectural Inverse Lighting

The premise of most architectural inverse lighting work is the elimination of design by iteration. Traditional lighting design involves an iterative process of configuring lights, rendering a result, evaluating the result, and then repeating until a desired goal is met. Inverse architectural lighting tries to eliminate this process by allowing a designer to code a goal so that a software system can find the light parameters needed to achieve that goal.

One of the first works in this area was [13] (see Fig. 1). This system found angles and intensities for lights given: 1) light positions manually specified by a user, and 2) numerically defined goals for the lighting design. Some of these goals were literal, such as goal irradiance to a surface while others were subjective, such as “spaciousness” and were derived from [14].

[15] used a similar approach to [13] and is the most recent and arguably most complete approach to inverse architectural lighting design. The work of [15] made only three assumptions about an environment: 1) surface reflectance can be described with symmetrical bidirectional reflectance distribution functions (SBRDFs), 2) the environment has no participating media, and 3) the used rendering model obeys the photon nature of light.

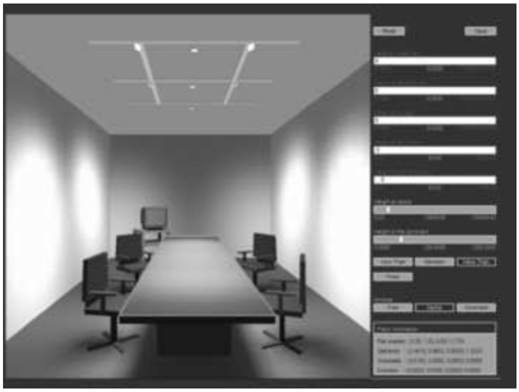


Fig. 1. An architectural space automatically lit by [13].

Within these assumptions, this system allowed designers to specify illuminations goals using scripts. For example, a designer might code a rule that constrained irradiance to a surface within a specific range. The system would then find light sources that fulfilled this and all other rules for a given space.

Inverse architectural lighting appears to be "solved". The work of [15] in particular produced a general purpose and flexible automated system for designing architectural spaces. This body of work would probably be useful to games for the design of "baked" lighting on in game architecture. However, the algorithms of inverse architectural lighting are too slow for real-time rendering and are effectively disjoint from the work of adaptive lighting tools for games.

4.2 Perceptual Inverse Lighting

Perceptual inverse lighting [16-19] tries to find the "best" solution to light an object, where "best" is defined using psychophysical theories and the goal of maximizing shape perception and detail. Objects are typically illuminated without user interaction.

The system of [18] found light source settings to maximize perceptual features such as shape, detail, and depth perception. [16, 19] targeted similar goals with different variations of a "light entropy" metric based on [20].

[17] is the most recent and complete work in this area. This work approached the problem differently from prior work. Objects were classified in the geometric domain into areas of local curvature. Illumination was then applied in a discontinuous fashion to each area to maximize goals per area. Cast shadows (see Fig. 2) were also explicitly considered, the only work in the field to do so.

Our work on adaptive lighting design tools is explicitly looking at interactive lighting, while the work of this field did not. However, similar to previous work, SAIL uses image analysis techniques, specifically, the light entropy metric of [16]. Unlike most of the work in this area, correlated geometry is not used. SAIL only needs an image to specify a goal or understand an object's appearance. Further, because analysis of images at runtime is too slow for real-time rendering, SAIL analyzes images as a preprocess and stores the results of this analysis for use at runtime.

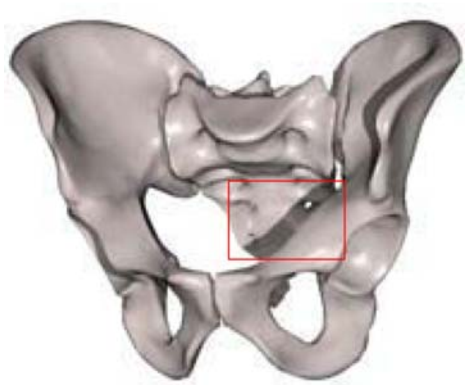


Fig. 2. The system of [17] has manipulated lighting to allow the indicated cast shadow to act as a depth cue.

4.3 Interactive Inverse Lighting

There are only two known works in this area: *LightKit* [21] and our earlier work the *Expressive Lighting Engine (ELE)* [8]. LightKit was a tool designed to interactively light medical models. It did not adapt to interactivity but simplified the iterative lighting process of traditional lighting design for non-experts.

5 Expressive Lighting Engine (ELE)

The Expressive Lighting Engine (ELE) is a dynamic, intelligent lighting system developed based on cinematic and theatric lighting design theories. It uses a constraint optimization algorithm to compose and adapt a lighting design dynamically and in real-time, accommodating user interaction while achieving artistic design goals. Design goals are input using numeric constraints that represent high-level artistic constraints that can adapt to changing dramatic and physical contexts.

ELE uses stage layout, scene graph information, and artistic constraints to create a light layout. ELE divides the scene into n different areas and then categorizes these areas as focus, non-focus, or background. This information is then used to achieve several artistic goals, such as directing attention to a focal point, increasing depth, and evoking mood by establishing visual contrast. To compute the light layout based on these goals, ELE minimizes a multi-objective function to determine the number of lights to use for each area:

$$p_{opt} = \arg\max_{p} \left(\lambda_v V(p) + \lambda_d D(p) + \lambda_m M(p) + \lambda_{vc} VC(p) \right),$$

where p is the light configuration and λ are weights representing artistic constraints. Specifically, λ_v is the importance of visibility, λ_d is the importance of depth, λ_m is the importance of modeling, and λ_{vc} is the importance of visual continuity. *V(p)* is visibility given p, *D(p)* is depth given p, *M(p)* is modeling given p, and *VC(p)* is visual continuity given p.

In determining the angles of light, ELE also takes into account the use of light angles in projecting depth, modeling, and mood, where mood is evoked through the angle of light on a character. For example, a character can be lit from below, creating a sense of evilness or mystery [10]. ELE uses nonlinear optimization to select an angle for each key light that minimizes the following function:

$$\lambda_v(1 - V(k,s)) + \lambda_- \left|k - k^-\right| + \lambda_m \left|k - m\right| + \lambda_l \min_i \left|k - l_i\right|,$$

where k and s are defined as the key light azimuth angle relative to the camera and the subject angle relative to the key light. k^- is the key light azimuth angle from the previous frame and λs represent artistic constraints. Specifically, λ_- is the cost of changing the key light angle over time (to enforce visual continuity), λ_m is the cost of deviation from the mood azimuth angle, m is the mood azimuth angle suggested by the artist, λ_l is the cost of azimuth angle deviation from a practical source direction, l_i is the azimuth angle of light emitted by the practical source i, and λ_v is the cost of deviation from an orientation of light that establishes best visibility.

The interaction of lighting colors in a scene composes the contrast and feeling of the entire image. Similar to the angle and layout systems, ELE uses a nonlinear optimization to search through a nine-dimensional space of RGB values. It differentiates among focus colors, non-focus colors, and background areas to select a color for each individual light in the scene. The multi-objective cost function evaluates the color against the lighting design goals, including establishing depth, conforming to color-style and constraints, paralleling dramatic tension, adhering to desired hue, saturation, lightness, and maintaining visual continuity.

While ELE received much success, it had several shortcomings. In particular, ELE was difficult to use due to the fact that the numerical constraints supplied were often interdependent in hard to understand ways. In addition, it did not consider the appearance of an object. SAIL was developed to overcome these shortcomings.

6 System for Automated Interactive Lighting (SAIL)

The *System for Automated Interactive Lighting (SAIL)* is an intelligent automated lighting system (see Fig. 3 and Fig. 4). Its design was based on an assumption that the appearance of an object is important when deciding how to light it. The current section presents the algorithms of SAIL. The next section presents a qualitative study of SAIL with lighting experts meant to confirm SAIL's contributions and contextualize automated lighting systems such as SAIL and the *Expressive Lighting Engine* (see Section 6) with lighting experts.

SAIL consists of two main components. The first component is image analysis, which it uses to understand a goal image. The goal image shown in Fig. 3 is a shaded circle, but the goal image can be of anything with shading and color contrast information. The second component of SAIL is object understanding, used to achieve adaptive runtime lighting. This component is very similar to the object lighting component of ELE with the addition of image analysis to understand how an object responds to changes in illumination. This component uses metrics extracted from a

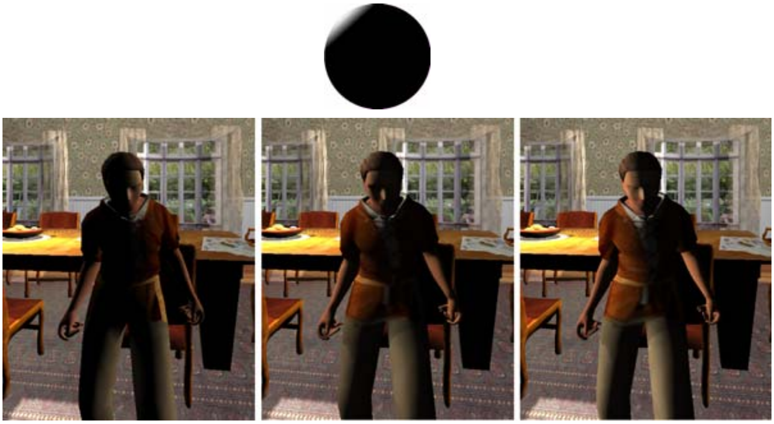


Fig. 3. This figure illustrates SAIL. The shaded circle is the goal image. The left image is the man naturally lit. The center image is the man lit to achieve the goal image. The right image is the man lit as a compromise between the two.

Fig. 4. The current implementation of SAIL.

goal image and the position of light sources in the environment. It then lights an object, using a three-point lighting model, to achieve an appearance between how the object "should" look based on the image and how it would look naturally based on the light sources in the scene.

6.1 Image Analysis

SAIL extracts two types of information from an image: 1) light direction, and 2) light contrast. Light direction roughly correlates to the direction of key light and light contrast roughly correlates to the intensity of fill light, but both are a combination of light direction and intensity and their effect on the object of an image.

6.1.1 Direction

Before further processing, each color channel of an image is filtered with a Gaussian kernel to remove high frequency information. This filtered image is then transformed into a grayscale image using the *Y* value from the sRGB derived *XYZ* color space as in [16]:

$$Y = 0.2126R' + 0.7152G' + 0.722B'$$

where *Y* is the *Y* term of the *XYZ* color space conversion from the sRGB color space and *R'*, *G'*, *B'* are defined by the function:

$$C' = \begin{cases} \left(\dfrac{C = 0.055}{1.055}\right)^{2.4} & C > 0.04045 \\ \dfrac{C}{12.92} & else \end{cases}$$

with *C* substituted for either *R*, *G*, or *B*.

Once the grayscale light image has been generated, two metrics for light direction are calculated. These roughly correspond to angles defined as shown in Fig. 5 but are actually derived from ratios of shading and pixel intensity in an image.

To find Θ as indicated in Fig. 5, local 3x3 windows of pixels are analyzed across the image, similarly to the application of a convolution kernel. For each window, the center of mass of pixel intensity is found relative to the window center. This relative position is then converted to a normalized 2D vector. The average 2D vector for the entire image is converted to an angle to find Θ.

Once Θ has been found, the term β as indicated in Fig. 5 is calculated by finding the average center of mass of pixel intensities along the axis *x'* as indicated in Fig. 5, specifically:

$$\beta = \frac{\sum_{i=1}^{n} \left(\dfrac{1}{\sum_{j=1}^{m_i} Y_{i,j}} \sum_{j=1}^{m_i} jY_{i,j} \right)}{n}$$

where *i* is a line out of n adjacent lines spread across all pixels in an image parallel with *x'* and perpendicular with *y'*, mi is the pixel count of line *i*, and Y_{ij} is the *Y* value of the j^{th} pixel of line *i*.

6.1.2 Contrast

Contrast is directly quantified by the light entropy metric of [16] but it is also a factor of the term β derived in the previous section. As β increases, contrast tends to increase (and entropy decreases) because of sharper shading gradients present in the image. This is why stochastic gradient descent is used at runtime to evaluate these terms as will be described in the next section. They do not correlate one-to-one with the parameters controlling light sources.

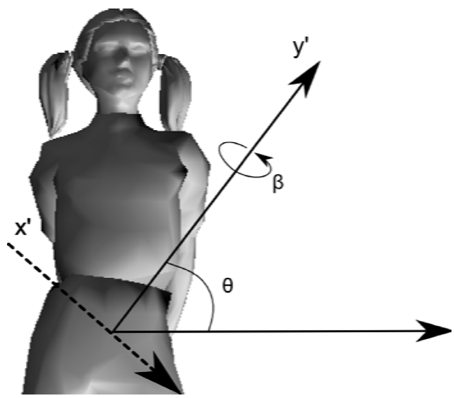


Fig. 5. This figure illustrates the approximate meaning of the two light shading direction extracted from an image.

We did not use the more recent light entropy formulation of [19] because our light direction terms provide information about the spatial configuration of pixels. The update to light entropy in [19] was primarily motivated by a desire to encode spatial information about an image and this is unnecessary in our case.

6.2 Modeling an Object's Response to Light

To adapt lighting at runtime, SAIL needs to understand how an object appears under different lighting conditions. Calculating this information directly from images of an object would be too slow to perform at runtime, so SAIL generates a model of how an object responds to light in a preprocess step.

The model is generated by jitter sampling the space of control parameters of SAIL's runtime three-point lighting model. This model consists of 5 control parameters: camera direction (specified as pitch and yaw), key direction (specified as yaw and roll), and key-to-fill ratio. For each randomly sampled control parameter set, the object is rendered. This rendered image is analyzed by the image analysis component and a set of image metrics (consisting of shading direction and contrast indicators as described in the previous sections) is produced. The surface of all randomly generated image metric sets is stored and used at runtime as a mathematical model of how an object's appearance changes in response to a change in illumination.

6.3 Runtime Lighting

Lighting is applied at runtime using a constrained set of lights based on the three-point lighting model. Three-point lighting consists of a key, a fill, and a back light. The key light provides dominant direction cues, the fill light reduces tonal contrast, and the back light rims the object, isolating it from the background. SAIL only uses a key and a fill light, as the effects of a backlight are high frequency and effectively removed in the Gaussian filtering step of image analysis. Capturing and applying the effect of a back light is left for future work.

The key light is a point light that is limited to only affecting the target object. The fill light is a directional light that is similarly limited. The control parameters for this configuration are a yaw and roll for the key direction and a key-to-fill ratio for the fill intensity. The fill direction is fixed at 90º relative to the key light. The desired key intensity is always 1. The actual key intensity at runtime is an average between the intensity as sampled from an object's environment and a value of 1.

SAIL understands the appearance of a model based on the jitter sampled surface derived as described in Section 7.2. It understands a goal image through image analysis. It derives the appearance of an object under natural lighting by evaluating the jittered sampled surface at the coordinates of the effective key and fill settings as derived from natural light source positions.

Lighting adaptation at runtime is achieved by applying stochastic gradient descent to move towards the point on the jitter sampled surface that is closest to the goal image and to the point that is closest to the natural lighting state. Closeness is measured using a least squares error metric between appearance metrics.

7 Study Design

This section details a qualitative study we conducted to explore the acceptance of SAIL to lighting design and visual design experts. Specifically, we discuss SAIL's usage of image as interface and SAIL's consideration of an object's environment to adapt lighting at runtime.

7.1 Participants

19 participants took part in the study. Participants were lighting "experts". An expert was defined as someone who had a personal or professional interest in lighting. This produced a range of participant expertise. Some participants described themselves as lighting designers and had a working vocabulary for the tools of light and a perspective that allowed them to purposefully control illumination to achieve a goal. Others were art directors or similar roles. These participants had an investment in light as a significant piece of their final visual work, but might not have the vocabulary or perspective to purposefully control light to achieve a goal.

Participants were recruited through personal contacts of the authors. Each was given a survey before a session started to gather basic demographic and expertise identity data. Most participants were male. All participants were 20 years or older. Age was fairly evenly distributed between 20 and 49 years of age inclusive. Fig. 6 shows the years of experience and specific areas of expertise of participants. Most participants had between 2 and 10 years of experience and a background in film and/or video games.

7.2 Study Design

Participants took part in 30-45 minute sessions conducted one-on-one with an investigator or a pair of investigators. Each session consisted of three parts.

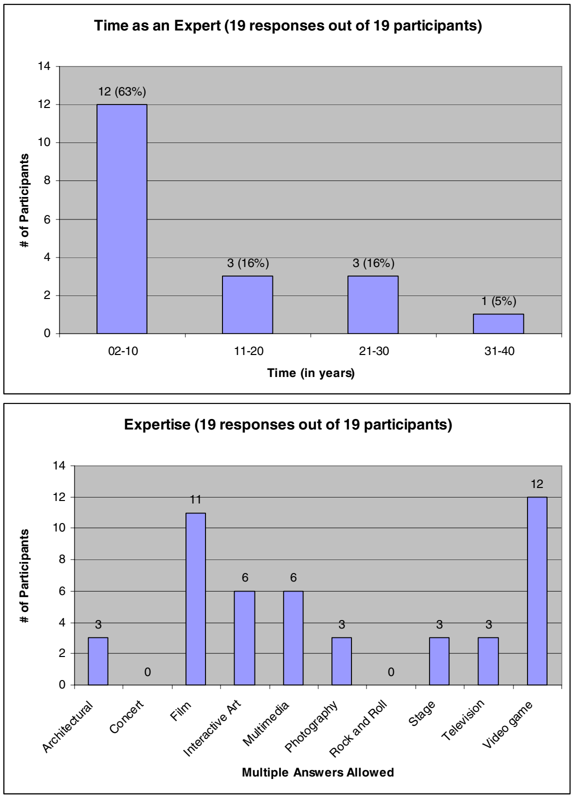


Fig. 6. This figure show participants' experience in years and specific areas of expertise. Participants were allowed to give multiple responses to the question regarding their area of expertise, so the individual counts do not sum to 19. Note that *Photography* replaces *Other* in the chart (*Other* was a fill-in choice on the survey). In every case where a participant wrote in an answer for their area of expertise they wrote *Photography*.

In part 1, participants interacted with the implementation of SAIL shown in Fig. 4. They could move the character, select between three goal images (one is displayed in the upper right of Fig. 4), select between three preset light states for the environment, and switch between "auto", "natural", and "attached" modes. "Auto" mode enabled SAIL, "natural" mode disabled SAIL and used only environment lighting, and "attached" mode enabled a "hacked" version of SAIL that completely disabled context considerations (SAIL applied lighting only in consideration of the goal image).

Participants were given enough context to allow them to understand the interface and the basics of SAIL. They were told that the character was being lit automatically and that the environment was not. They were not given specific information about SAIL's algorithms or intended purpose. Participants were asked to interact with the environment until they felt that they could discuss the aesthetics and function of lighting in the demo and particularly, the behavior of lighting on the character.

In part 2, participants were interviewed using a semi-structured approach. A set of questions was used to guide discussion but individual questions could be skipped or additional questions asked based on conversation flow. Participants were not given specific information about SAIL and the investigator deferred answering these questions. For example, if asked, “I’d like to know what you’re trying to accomplish here” the investigator might respond, “Let’s get back to that later.”

In part 3, participants were given complete information regarding SAIL’s algorithms and anticipated applications. Any questions such as, “What lighting model are you using?” or “I’d like to know what you’re trying to accomplish here” were answered at this time.

We approached the design of this study from a deductive, qualitative mindset. We wanted data to argue evidentially about SAIL’s usage of images as input and SAIL’s consideration of appearance to adapt lighting at runtime. We felt that the views and interpretations of lighting experts were important, because lighting experts are the target of our work and it is their expert perspective that determines whether SAIL is a contribution or not.

Giving participants incomplete knowledge during parts 1 and 2 was intended to encourage them to make comments at a “mid level”, i.e. comments that would be specific but not so specific as to have little meaning outside the very particular constraints of the demo. Further, it was our hope that it would encourage them to discuss their process. We feel this was successful. For example, when asked what their interpretation of the image was, many participants explained their interpretation and also explained why they interpreted it as such.

Part 3 was included for participant satisfaction, as it was expected many would have specific questions after parts 1 and 2 about SAIL (many did). Further, it offered an opportunity to collect directed suggestions from participants about SAIL and lighting design tools in general.

Passages from transcripts created of part 2 were categorized as evidence for either SAIL’s usage of images or as evidence for SAIL’s consideration of an object’s environment at runtime. This data was then further coded into common themes as they were identified, which are discussed in the next section.

8 Study Results

SAIL makes two contributions to the design of lighting tools. First, it presents a lighting design tool that uses images as an interface to specify lighting goals, which we believe is understandable for designers. Second, it presents a lighting design tool that adapts the lighting to reflect the natural lighting of an environment as well as keep the desired outcome, which we believe is a desired feature of a lighting system. Note that the numerical codes used to identify participants throughout the rest of this chapter can be used to find the complete transcription of a participant’s session in [7].

8.1 Images as Interface

Our study indicates that the semantics of the images used in the demo were understandable to participants, and thus could lead to a usable interface for a lighting

Fig. 7. This figure illustrates the three images available to participants during evaluation. The leftmost image is the image attached to key “Q”. The center image is the image attached to key “W”. The right image is the image attached to key “E”.

design tool. The three possible images a participant could select from are shown in Fig. 7. Images represented a “lit sphere”, a technique used to represent shading in illustration [22]. Each image represented lighting direction and light contrast. The meaning of the image as light direction was clearly understood by participants, although some were confused about its indication of light contrast. Other participants were confused by deviation between what the image displayed and what they saw on the character while SAIL was considering context. Specifically, the presence of a fill light might imply the elimination of absolute darkness, so the presence of absolute darkness in an example image can be confusing. Further, deviation between what a displayed image indicates (the ideal goal) and what SAIL is doing at runtime (a compromise between the ideal and the natural) can be confusing. Essentially, as described by one participant, if the lighting in the scene can deviate from what is indicated by the image, then the logic of what the image represents is violated.

SAIL interprets an image as visual cues indicating light direction and contrast. Participant 10_27_01 appeared to interpret the image in the same manner:

INVESTIGATOR:
So, is there any sense what it’s doing specifically right now?
10_27_01:
Well, when I look at the image and I look at the character I can see it does seem to be the direction of the lighting and the overall shading. I’d be curious to see how much I could mess with that image.

As did participant 10_28_04:

INVESTIGATOR:
What do you think the role of the image is? The image in the upper right.
10_28_04:
The lighting on the character. Or, if that was her face, if her face was a sphere, that’s how it would look.

And participant 10_28_02:

INVESTIGATOR:
Go into a little bit [pause] What’s your interpretation of what those images mean? In terms of what the system’s doing with them?
10_28_02:
My interpretation is I’m looking at the source of the lighting. When I look at this [referring to image attached to the key Q] I see there’s a top 45-degree angle [pause] It’s almost a rim light. It’s fairly far back, so it’s not lighting the front. It’s not all the way back ‘cause I can still see, it’s a highlight. This [referring to the image attached to key W] I’m seeing a light with more spread, lighting from the

bottom-right quadrant, but a fairly hard light, in the sense that this is in shadow [10_28_02 points out the upper-left corner of the image]. It's not diffused enough to get me bounce back here. This [10_28_02 points to image attached to key E] I'm looking at the same idea but from more [pause] more top-ie. It seems a little more top-ie even than if you were to switch this one that I'm looking at [10_28_02 indicates the image attached to key Q]. [pause]. [It shows] Where's the light coming from. If this were the key light, [it shows] where's the source and what's the nature of the source.

Participant 10_30_02 explains why an image of a shaded circle might make sense to a lighting expert, comparing it to an "environment map":

INVESTIGATOR:
What's your interpretation of the image?
10_30_02:
What I get from the image is like in an environment map, it's maybe the translation I would do. It's like a map that you would multiply on top of the whole character. For example, if you [pause] What I was looking for was if this [10_30_02 is pointing on image #2] has a dark area on the top, I will get that always on the top of the character and it's going to be brighter going down. That's going to be multiplied to the frame buffer on the character area. That's what I interpreted.

And participant 10_30_03 made a similar comparison to "hemispherical lighting":

INVESTIGATOR:
What's your interpretation of the image?
10_30_03:
Of the…
INVESTIGATOR:
Ya, the thing that's surrounded in the magenta color.
10_30_03:
I would expect that would be a spherical representation of what the lighting's doing. So a hemisphere lighting solution [pause] Not lighting solution but if I was viewing hemisphere lighting it would be in a similar format. So it's basically where the lighting is coming from if you were looking at it on a globe.

All other participants described the image as indicating direction. Some participants gave responses that were unclear with regards to whether the image indicated light contrast or not:

INVESTIGATOR:
What do you think is the role of the image? The thing in the upper right corner.
10_28_03:
Oh, I have to admit [pause] I only read it as the source of the primary light. Might have meant something different but I have to admit I didn't…[pause]

And participant 10_26_01 did not think the image represented a fill light, which controls contrast:

INVESTIGATOR:
On that note, what do you think the role of the image was then?
10_26_01:
The role of the image seemed like a quick, top-down sort of chart of kind of where the key light was coming from. I didn't get much feel for the fill light in this. […]

Confusion regarding a fill light is potentially described by a similar confusion and comment from participant 10_29_02…

INVESTIGATOR:
Ok. What do you think the role of the image is? In the upper right?
10_29_02:
It's a model of the light setting parameters that you're using [pause] Under that setting. So it's just a representation, I think, of the direction of light and the quality of light. So that's a smaller light source [10_29_02 has the image attached to Q selected] and it models the direction it's coming from. Possibly without any fill light [pause] [10_29_02 interacts with the software system].

…and additional conversation during the freeform conversation of part 3. Participant 10_29_02 further explained the comment "Possibly without any fill light". To participant 10_29_02, a fill light in a film three-point lighting model is used to set a minimum contrast and eliminate any absolute darkness. However, all of the images that participants could select from had an absolutely dark portion (see Fig. 7). SAIL was trying to recreate this feature and left part of the character completely dark. This confused participant 10_29_02 and resulted in doubt about whether there was a fill light at all.

Participant 10_26_02 was confused about the role of the image in general but still described its semantics as both indicating direction and contrast:

INVESTIGATOR:
You mentioned the image in the upper right corner, could you go into a little bit more detail about what you think the role of the image is?
10_26_02:
I actually don't really know…
INVESTIGATOR:
Do you have any sort of inkling or would you venture a guess?
10_26_02:
I assume it's where the lighting is coming from so, this [indicating the image in the software system] would be where the light is coming from the top or directionality as opposed to [long pause, while changing the image in the demo]. Ya, I'm not quite sure. If I had to venture a guess I would probably say the intensity and the directionality of the character lighting. But I didn't see it on her so I wasn't sure if I was right or not.

Participant 10_29_01 had similar confusion:

INVESTIGATOR:
Related to that, what do you think the role of the image is, in the upper right?
10_29_01:
I would imagine that it's sort of the lighting cast on the character. That's what I'm guessing. [pause] By the individual light sources. I'm not entirely sure, actually, to be honest, now.
INVESTIGATOR:
You're in auto mode, correct?
10_29_01:
Let me put it in auto mode…
INVESTIGATOR:
Try moving her a little bit after you change the image.
10_29_01:
To be honest I still don't understand it.

This confusion is possibly explained by participant 10_30_05:

INVESTIGATOR:
What is you interpretation of the image in the upper right?
10_30_05:
I'm not really sure what you're trying to indicate there [pause] If it's a direction of light source because it can be connected and can be disconnected from what the lighting in your scene is, then it breaks the logic for what it's trying to tell me. I think. I got confused when I was switching back and forth between the different modes as to what this was actually trying to do. I guess [pause] I'm familiar with lighting things in applications that have just a sphere to tell you where the light source is [pause] I'm kind of used to seeing these kinds of things and I would expect that to be an uplight, kind of very broad. Key light, rim light.

Similar confusion from participant 10_27_01 was alleviated when SAIL's "auto" mode (the mode in which SAIL considers natural lighting) was disabled and SAIL was put into "attached" mode (where SAIL only recreated the lighting of the image):

10_27_01:
So, looking at the image is my guess is it's a simplified lighting model. That allows you to use an image to basically light the scene or, like you said, just modify what's going on in there, hopefully there's some influence from something else. I could see something like that being used possibly on the character but certainly not on the backgrounds at all, that wouldn't work very well. If I was going to use a lighting model like that I would probably restrict it to the character and do the backgrounds a different way.
INVESTIGATOR:
So, is there any sense what it's doing specifically right now?
10_27_01:
Well, when I look at the image and I look at the character I can see it does seem to be the direction of the lighting and the overall shading. I'd be curious to see how much I could mess with that image.
INVESTIGATOR:
So actually manipulate the image, like in Photoshop or something?
10_27_01:
Ya, to see if it's strictly that image that's driving the lighting or if it's a multiplier of it or whatever. It'd be interesting to see if I could get something a little more realistic out of it. Overall, [pause]. So this one here doesn't respond like I'd think though. I've got this image here that's just really bright in the upper left corner and the rest of it's dark. But I don't seem to be seeing that on the character. So maybe I'm interpreting that wrong.
INVESTIGATOR:
Well, this system does take into account both the system and the environment. So, if you [pause] The attached mode. That is purely the image.
10_27_01:
Ok ya, that's more what I'd expect to see.

Overall, an image of a shaded sphere appears to be an understandable interface to lighting experts. Confusion seemed to stem from how SAIL was applying this image in auto mode. An image for specifying lighting goals appears to be a reasonable approach for future work, but care must be taken when presenting this image to a user. Specifically, a fill light has inherent semantics that should be considered (a fill light potentially implies removal of absolute darkness). Also, it can be confusing if

Fig. 8. Many participants expected the windows in the back of this room to be sources of illumination even though they were not.

the image only reflects the desired lighting goal when SAIL is applying lighting to achieve a compromise between the desired and natural goals.

8.2 Reflection of Natural Lighting

Based on the data collected, SAIL is successful at reflecting natural lighting while applying a lighting goal at runtime. However, many participants felt that the natural lighting state itself was bad. SAIL's successful reflection of natural lighting was indicated mostly by comments from participants who identified bad lighting that was present in both natural and auto modes.

This section includes discussion of what experts found unsatisfactory about the environment in addition to evidence for SAIL's success at reflecting natural lighting. This discussion is included as indication of what future work on SAIL should possibly consider to produce a tool that is not only predictable but also satisfying in its results.

This section is divided into two parts: 1) lighting was not sufficiently motivated by logical light sources in the environment, and 2) exceptions to the rule are necessary tools in the toolbox of lighting experts.

8.2.1 Logically Motivated Light Sources

Lighting direction was not sufficiently motivated by logical light sources. The first room (see Fig. 8) contained two large windows, picturing an outside spring day. Logically, these should be large sources of illumination. But they were not, as noticed by participant 10_29_02:

> **INVESTIGATOR:**
> What are your general impressions regarding this system?
> **10_29_02:**
> It's a simple system. There's a few things I have questions about. For example, there are windows over here. I'm just wondering why there are no lights that represent the windows, unless [pause] 'Cause I don't see a change going from natural to… [pause]

INVESTIGATOR:
Ya, there's no lights representing the windows.
10_29_02:
So, I would say your ambient lighting, at the very least, should account for actual sources of light, and not just aesthetic sources. When I switch to setting 2 [referring to the environment preset attached to key 2], there seems to be a light source here [10_29_02 indicates the light source in the ceiling about the dining room table in the first room] but it doesn't seem to cast on the table. That should be there. Also, it doesn't affect the character in an expected way for me. This [10_29_02 indicates the lighting in the room with the dining room table] looks like it has direction like a directional light but [pause] As this [indicating the point of light in the ceiling above the dining room table] is the only point source that I can see in the room, yet it's [referring to the lighting of the character] coming from almost directly above. Just wondering what that choice means [note: 10_29_02 has image Q selected during this dialogue].

Participant 10_26_01 found the lighting in general to be "non-diegetic":

INVESTIGATOR:
Would you describe the lighting as satisfactory in the environment?
10_26_01:
As I said, the fact that the lighting seemed non-diegetic to me [referring to a comment made by 10_26_01 while 10_26_01 was interacting with the system before the interview portion of the session] was a little bit confusing.
INVESTIGATOR:
Could you explain what you mean by "non-diegetic"?
10_26_01:
Sure. And diegetic might not even be used frequently with lighting it's used with sound all the time and film and that means that there's a difference between when you have the soundtrack playing in the background and when you have the sound seem to be coming from the radio in the car that the person is sitting in. That's a diegetic sound, a sound that is actually emanating from the environment in some way rather than sound. So, I guess the real term, the stage term would be a practical light. You can have a lamp sitting on stage that is actually casting light or you can have a lamp sitting on stage and have a light that is up in the rafters that is trying to simulate what that light would look like coming from the lamp.
And so the one thing that I found not satisfying about the lighting in the scene as I mentioned [referring to a comment made by 10_26_01 while 10_26_01 was interacting with the system before the interview portion of the session] was the two lamps didn't seem to be casting the light. They seemed to be representing light sources but they were representing lighting sources that were outside the room. That was a little by confusing to me.
INVESTIGATOR:
By the two lamps, you mean the lights indicated by the arrows?
10_26_01:
No, actually what I mean is the physical lamps sitting on the chest of drawers here [10_26_01 points to a lamp in the 3D scene that is located on a chest of drawers in the center room of the environment] and the one sitting on the table in the third room.
INVESTIGATOR:
So would you describe this more as a problem with the preset environment lighting?
10_26_01:
Yes, absolutely. […]

The lack of lighting motivated by logical light sources resulted in an environment that overall was considered too "flat":

INVESTIGATOR:
So on that note, what are your general thoughts regarding the behavior of the system?
10_27_01:
My general thoughts are that it doesn't look like much has been put into the background at this point. I like to see a lighting environment that really looks like it all works together. I think part of that is I'm just really only evaluating the lighting on the character but it's kind of hard to tell [pause] The character doesn't fit in with the environment so there's no relationship…
INVESTIGATOR:
Could you go a bit more into [pause] How specifically? [referring to the 10_27_01's statement that the character does not fit into the environment]
10_27_01:
Specifically the environment looks like your typical evenly lit no-lighting environment. I can't see any distinguishing shading other than a very kind of high-contrast from the key light. There's not much fill there, there's no shadows. There's no ambient occlusion in the corners or anything like that it's just basically really flat.

SAIL appeared to successfully reflect the flatness of the environment when lighting the character. Participant 10_28_02 described SAIL's "auto" mode as reflecting the "kick" off the environment while the "attached" mode (where SAIL completely ignored the environment) only reflected the "global source", where "global source" was participant 10_28_02's description of light direction and contrast as indicated by the images in the demo:

INVESTIGATOR:
Go into a little bit [pause] What's your interpretation of what those images mean? In terms of what the system's doing with them?
10_28_02:
My interpretation is I'm looking at the source of the lighting. When I look at this [referring to image attached to the key Q] I see there's a top 45-degree angle [pause] It's almost a rim light. It's fairly far back, o it's not lighting the front. It's not all the way back 'cause I can still see, it's a highlight. This [referring to the image attached to key W] I'm seeing a light with more spread, lighting from the bottom-right quadrant, but a fairly hard light, in the sense that this is in shadow [10_28_02 points out the upper-left corner of the image]. It's not diffused enough to get me bounce back here. This [10_28_02 points to image attached to key E] I'm looking at the same idea but from more [pause] more top-ie. It seems a little more top-ie even than if you were to switch this one that I'm looking at [10_28_02 indicates the image attached to key Q]. [pause]. [It shows] Where's the light coming from. If this were the key light, [it shows] where's the source and what's the nature of the source.
INVESTIGATOR:
Would you say based on your interpretation that what the system is doing makes sense?
10_28_02:
Yes. I'd want to spend a bit more time with it to see how wandering around in this environment, what the implications were? 'Cause this room for example [indicating the first room with the dining room table], there's more kick coming off the walls. So this [referring to the auto mode] gives me sort of my ambient idea and then the attached [referring to attached mode] looks like it's just being lit from the global source.

INVESTIGATOR:
By global source, you mean the image right?
10_28_02:
Ya, these guys [10_28_02 indicates the images displayed in the system]. So it's independent of …[pause]

Participant 10_27_01 described the lighting in natural and auto modes as comparably poor:

INVESTIGATOR:
So the environment seems very flat, but she [referring to the character in the software system] doesn't, or…?
10_27_01:
She seems too contrasty. It looks like there's a bit of a fill light there, in some cases I can see a bit of a rim light. In one of those modes, little bit. But overall it's just way too harsh. It's not adapting well…what I want to see here is certainly more of a fill light so that I can at least see her features [note: during this sentence, the 10_27_01 is pointing out the character being lit under modes that result in significantly portions of her body being completely black or nearly completely black]. Usually what'll happen in an animated series and such is that they'll have a separate pass. The background's gotta look good and the character's gotta look good. And they'll often have a rig that even travels along with the character, so they'll always be a rim light to a certain amount. And they'll just kind of cover those basic three-point lighting principles.
INVESTIGATOR:
So is that notably about the same in natural mode, 'cause currently you're in auto mode so if you switch it to natural mode?
10_27_01:
The natural mode is better. But it's still [pause]. Some of that may just be the gamma of this screen too, 'cause it looks a little bit washed out. I guess the thing that I noticed the most is that there doesn't seem to be much interaction with the environment. And, well here I can see that I'm passing underneath a light source. A lot of this could be not necessarily the technology but just how the lights were placed. I think you can work with whatever you got, and it's just a matter of placing the lights a little better.
Participant 10_30_01 noted an unmotivated light source in both natural and auto modes:
INVESTIGATOR:
So you would say that in natural mode, the lighting between the character and the environment doesn't match? Or what do you mean by that?
10_30_01:
I don't know what kind of technique you use but sometimes when you walk through the rooms of the environment you don't really see the lighting change that much. Just the interaction between the lighting on the character to the environment, it's not that noticeable. For instance, you can see here…
INVESTIGATOR:
So the lamp in the center room?
10_30_01:
Ya. You see here that you obviously have a dominant light source coming from this side but your character is not lit from behind [behind being the direction towards the light source in this case]. [pause] [participant interacts with demo] It would be interesting, since now you have the ability to mix between your preset lighting and your environment lighting, it would be interesting if you could set the scene, to set

multiple lighting sources, if you can. Maybe one is from in this area maybe if you want to say the light is on you can have your light source put in there to light a character and then another in here, maybe you can set another light source, just another coming from in there or have some lights from this angle to simulate that lighting. And you can just lerp [linear interpolate] between them to get more interactive [pause] You'll pick up the environment better.
INVESTIGATOR:
Is that all specifically the natural mode or is…?
10_30_01:
No, it's not just natural mode, it's when you have the auto mode which is kind of a blend between natural and your attached lighting. I'm just saying, for your attached to set up multiple light sources, you can use your attached lighting to do that. I don't know exactly how you do this. But I would say just in general since you have this ability to blend [pause] You have a natural and an attached, you can use an attached lighting mode to set up some multiple light sources and have your character interact, it would be more interesting.

And participant 10_26_02 noted a general mismatch between lighting on the character and lighting on the environment that was comparable in both auto and natural modes:

10_26_02:
And I think that might be an "attached" lighting issue. I always feel like players feel pretty flat when they don't have shadows. And most of these rooms feel really, really over lit compared to her. 'Cause she's got …
INVESTIGATOR:
So you would say the rooms seem brighter than she is?
10_26_02:
Ya. She's reacting really nicely to the light but I don't think that it reflects it in the world very well.
INVESTIGATOR:
Is that specifically in a specific mode?
10_26_02:
It's more the mismatch between the way she's lit [pause]. When I expect her to have black on the back and the lighting kind of striking her like this I expect it to be darker on the back end. I'm assuming that's because it's vertex lighting and not many…
INVESTIGATOR:
But specifically, that seems to be the case in any mode?
10_26_02:
Mm-hmm. She tends to be lit pretty well. But the world isn't really matching the way she's being lit. But I was looking pretty closely at just her lighting rather than the world so I didn't really notice until now.

SAIL was apparently successful at reflecting natural lighting, but the lighting itself was considered poor by most participants. This specifically may not be a problem to solve. In contexts where SAIL might be useful, we can expect natural lighting to be designed in a satisfactory way. Therefore, SAIL would have better data. However, for applications where better data is not an option (i.e. SAIL integrated into an educational environment and used as a tool to make lighting design easier), SAIL could possibly be integrated with our previous work the *Expressive Lighting Engine (ELE)* [8] to automatically illuminate the environment.

8.2.2 Exceptions to the Rule

Our study indicates that any lighting design tool should allow designers to break rules. The previous section argued that lighting motivated by logical sources of light is important. This can be identified as a rule that an automated system such as SAIL should follow. However, what is an equally important "rule" is a lighting expert's ability to break rules if desired. Both participants 10_29_03 and 10_28_04 found "attached" mode to be interesting for its ability to detach the character lighting completely from the environment and therefore completely ignore logical sources of light:

> **10_29_03:**
> What I really am most fond of is this ability to detach [pause] I like that ability to detach the figure from the environment. Just because it's a kind of shadow like effect where I can make things that I regard as less important [pause] But I sure would like to be able to light something in the background. Oh, I don't know why because in the context of a game maybe you don't want to do that, there's certainly the game isn't going to be about designing the lighting.

And participant 10_28_04:

> **INVESTIGATOR:**
> What are your general reactions to the behavior of this system?
> **10_28_04:**
> I thought it was neat how you could detach the lights [pause] Like, this is a scary look [participant has selected image attached to letter W] with this, but if you put this it doesn't quite match.
> **INVESTIGATOR:**
> So you're talking about attached mode, specifically?
> **10_28_04:**
> Ya, attached mode. It gives a more cinematic look I guess. If this were a game you could create more cinematic looks on [pause] Really highlight your character how you want. Independent of the environment or [pause] You just have more control I guess. I haven't done any lighting in games but [pause] Ya, it sounds like you'd have more control.

This was echoed by participant 10_28_03 who described it as "my favorite mode":

> **INVESTIGATOR:**
> So, overall would you call the lighting satisfactory?
> **10_28_03:**
> In some modes in some locations. The night location was in some ways the most interesting but only if there was a source of strong light. If she was on the other side of the room or further down this way [participant indicating the corner of the room with the dining room table near the entrance to the adjacent room] there just wasn't any light coming in on her, and the night mode [pause] Or if she gets over here [participant puts the character completely in the corner between the room with the dining room table and the room with the radio] she's just completely dark.
> **INVESTIGATOR:**
> Well, actually that could be attached mode [pause] I think you're in attached mode. Attached is meant to be intentionally sort of very, very disconnected from… [pause]

10_28_03:
If I'm in auto there [pause] The night in auto there, you can't really do very much [pause] And then in attached mode it becomes quite interesting in fact that may be my favorite mode, in truth.

Also, despite the general dislike of complete darkness, participant 10_29_06 described a possible usage for this effect:

INVESTIGATOR:
What are your general reactions to this?
10_29_06:
I'm not sure what's going on [pause] Why the face is so dark while I'm in this room. 'Cause when I went to the other rooms and hid behind a wall, the same reaction did not happen. It seemed that it did not take the situation of blocking. I also wanted to see if it's going to make me feel different when I was changing different light modes. Yes and no. The person didn't really react [pause] I couldn't really tell between E and W [referring to lighting images attached to E and W] but in this mode, which one is it? In the auto one, it seemed to be the better looking one, although again you have these shadows. In advertising you would never want to do that unless [pause] You don't want them o become so dark unless there's a reason for it. I was trying to figure out what the reason was she had a shadow cast on her in this spot when [pause] I understand that there's a light up here but it just seems strange that when you're closest that you wouldn't have a corrective light to shine on her face. The one that I liked the most was the auto mode, that wasn't too dark.
[…]
INVESTIGATOR:
Specifically, would you describe the lighting as satisfactory? You talked a little bit about it but…
10_29_06:
It depends on what exactly you want to do. If she was trying to move [pause] I'm a pensitive person and I come here and now I actually realize that I'm guilty [10_29_06 has moved the character to a spot where much of her face is complete blackness] where she slowly transitions to dark, then I get it. But to me it's a bit dramatic in this case for no reason. I don't get the reason, I don't see the context. This one seems too plain to me, the S, the natural one. She seems very flat. [pause] But I'm not getting the same effect. Like, I don't understand why this effect happens here. It's funny how you're trying to see the details of her face and you can't. I think that kind of bothers you from the beginning because I think people have a tendency to really focus on facial features.
Ya, it becomes too dark too soon, I guess.
The effect of turning off the light it feels to me like the end of Act I or something like that [10_29_06 has transitioned room lighting cue to cue 2, the "nighttime" cue].

Our conclusion is that experts may want to completely violate any rule that is inherent to an automated lighting system. Although participants were able to detach lighting using the "attached" mode of the demo, SAIL was never designed with this mode in mind. It was effectively a "hack" for comparison purposes. SAIL has no manner in which to react or accommodate an exception such as completely detached lighting but evaluation indicates that these exceptions are desired and will be used in any practical application of SAIL. Therefore, future work on SAIL should incorporate exceptions in an integrated fashion.

8.3 Conclusions – A Look into the Future of Lighting Design Tools

This chapter presented discussion on lighting design tools, specifically discussing the *Expressive Lighting Engine (ELE)* and the *System for Automated Interactive Lighting (SAIL)*, and a study showing some lessons learned about what visual designers would expect from their lighting design tools. Our study indicates several important implications for future lighting design systems and tools for interactive entertainment. Specifically, we conclude that (1) images can be an effective interface for lighting experts to specify lighting goals, (2) while an automated lighting system needs to consider context, the success or failure of the lighting appearance is really a matter of how well designed the context is, and (3) any lighting design tool should allow designers to break the rules under which the automated system operates.

ELE and SAIL represent a new transition towards a middleware system for visual design, specifically lighting design, within interactive entertainment. Breaking this ground is important yet it is a journey that is full of exciting challenges within the design, business, and technical areas. In 2004 during a talk at Penn State University, John Buchanan, formally University Research Liaison at *Electronic Arts* (world leading game company) and currently the Technology Director at Relic Entertainment, stated that the game industry is moving towards being a content development industry rather than a software industry. By that he meant that the game industry is focusing more on development of content rather than creating systems or tools. To accommodate this content focused industry. many middleware companies are starting to appear, such as *Simutronics (Hero Engine)*, *Epic Games (Unreal Engine 3)*, *Terathon Software (C4)*, *PathEngine (PathEngine)*, *Havok (Havok Physics)*, *Ageia (PhysX), Avid Technology (Alienbrain), Emergent Game Technologies (Gamebryo), IDV, Inc. (SpeedTree),* and others. Clearly, what was once a niche industry populated mostly by physics middleware has now grown into a diverse and essential complement to game content development.

While some lighting middleware does exist (*Synapse Gaming's SunBurn Studio, Illuminate Labs's Beast,* and *The Game Creators's giles)*, it is primarily targeted at quickly generating high quality "baked" lighting. Tackling the challenging design problems exclusive to video games has remained mostly unexplored. This chapter has scratched the surface of such challenges; we anticipate this direction to grow as it offers many exciting opportunities for developers of interactive entertainment.

Acknowledgements

Special thanks to David Milam for modeling the environment and modeling and animating the character used in the demo (http://www.sfu.ca/~dma35/aboutme.html).

Keywords

Bidirectional Reflectance Distribution Function (BRDF) – A bidirectional reflectance distribution function (BRDF) is a mathematical function that describes how a surface reflects illumination. It maps incoming illumination to outgoing illumination.

Entropy – The entropy referred to in this chapter is information entropy. Entropy is a quantification of the information potential of a variable.

Indirect Light – Indirect light is also called "bounce" light. Indirect light is any light that is emitted into an environment from surfaces that reflect the primary sources of illumination. For example, a light bulb is typically a form of direct light while a tabletop near a light bulb is a form of indirect illumination.

Irradiance – Irradiance is a unit of measurement describing the amount of light incoming or incident to a particular area.

Lightmap – Lightmaps or "baked" lighting refer to a computer visualization technique. Light is precalculated and stored and then later applied at runtime in real-time rendering.

Radiosity – Radiosity is a lighting algorithm that captures important light phenomenon such as bounce light to create soft, believable lighting. Radiosity is computationally expensive but the generated data of radiosity is static so long as an environment remains static. As a result, radiosity can be calculated in a preprocess step, saved, and then later applied in a video game at runtime, making it ideal for producing lightmaps.

Three-point Lighting – Three-point lighting is a lighting technique that divides lighting usage into three goals: key lighting to control direction cues, fill lighting to control light contrast, and back lighting to control an object's silhouette and separation from the background.

Resources

Articles

Patow, G. and X. Pueyo (2003). "A Survey of Inverse Rendering Problems." COMPUTER GRAPHICS forum 22(4): 663-687.

Patow, G. and X. Pueyo (2005). "A Survey of Inverse Surface Design From Light Transport Behvaior Specification." COMPUTER GRAPHICS forum 24(4): 15.

Books

Akenine-Möller, T. and E. Haines (2002). Real-Time Rendering. Natick, Massachusetts, A.K. Peters, Ltd.

Birn, J. (2006). [digital] Lighting & Rendering. Berkeley, CA, New Riders.

Carter, B. (2004). The Game Asset Pipeline. Hingham, Massachusetts, Charles River Media, Inc.

Millerson, G. (1991). Lighting for Television & Film. Linacre House, Jordan Hill, Oxford, UK, Focal Press.

Zettl, H. (1999). Sight Sound Motion: Applied Media Aesthetics. Belmont, California, Wadsworth Publishing Company.

Conferences

ACMSIGGRAPH – Association for Computing Machinery's Special Interest Group on Graphics and Interactive Techniques

DiGRA – Digital Games Research Association

EUROGRAPHICS – European Association for Computer Graphics
GDC – Game Developers Conference
GLS – games + learning + society

Journals

ACM Computers in Entertainment
Game Studies, the International Journal of Computer Game Research
United Business Media Game Developer Magazine

References

ESA, Essential Facts about the Computer and Video Game Industry, ESA. p. 16 (2008)

ESA. Video Games & Families (2008), http://www.theesa.com/gamesindailylife/families.asp (cited 2009 January 8, 2009)

Hyderabad, C.: FXLabs, Eros Intl launch 3D game on Ghanjini, in Business Standard (2008)

Block, B.: The Visual Story: Seeing the Structure of Film, TV, and New Media. Focal Press, New York (2001)

Birn, J.: [digital] Lighting & Rendering, 2nd edn., p. 416. New Riders, Berkeley (2006)

McTaggart, G.: Half-Life 2 / Valve Source Shading, Valve Corporation. p. 97 (2004)

Zupko, J.: A System for Automated Interactive Lighting (SAIL). The Pennsylvania State University, State College, Pennsylvania (2009)

Seif El-Nasr, M.: Automatic Expressive Lighting for Interactive Scenes, Computer Science, p. 147. Northwestern University, Evanston (2003)

Callahan, S.: Storytelling Through Lighting. In: SIGGRAPH. ACM, New York (1996)

Millerson, G.: Lighting for Television & Film, 3rd edn., p. 466. Focal Press, Oxford (1991)

Reid, F.: The Stage Lighting Handbook, 6th edn., p. 217. Routledge, New York (2001)

Patow, G., Pueyo, X.: A Survey of Inverse Rendering Problems. COMPUTER GRAPHICS forum 22(4), 663–687 (2003)

Kawai, J.K., Painter, J.S., Cohen, M.F.: Radioptimization - Goal Based Rendering. In: ACM SIGGRAPH 1993 (1993)

Flynn, J.E., A study of subjective responses to low energy and nonuniform lighting systems. Lighting Design and Application (1977)

Costa, A.C., Sousa, A.A., Ferreira, F.N.: Lighting Design: A Goal Based Approach Using Optimisation. In: EUROGRAPHICS (1999)

Gumhold, S.: Maximum Entropy Light Source Placement. In: IEEE Visualization 2002, pp. 275–282 (2002)

Lee, C.H., Hao, X., Varshney, A.: Geometry-Dependent Lighting. IEEE Transactions on Visualization and Computer Graphics 12(2), 197–207 (2006)

Shacked, R.: Automatic Lighting Design using a Perceptual Quality Metric. Computer Science and Engineering, p. 68. The Hebrew University of Jerusalem, Jerusalem (2001)

Vázquez, P.-P., Sbert, M.: Perception-Based Illumination Information Measurement and Light Source Placement. In: ICCSA 2003 (2002)

Shannon, C.E.: A Mathematical Theory of Communication. Bell System Technical Journal 27(1), 379–423, 623–656 (1948)

Halle, M., Meng, J.: LightKit: A lighting system for effective visualization. In: IEEE Visualization 2003, pp. 363–370 (2003)

Sloan, P.-p.J., et al., The Lit Sphere: A Model for Capturing NPR Shading from Art, Microsoft Research (2000)

Chapter 5

Utilization of Fuzzy Rules in Computer Character Animation

Piotr Szczuko and Bożena Kostek

Multimedia Systems Department, Electronics, Telecommunication and Informatics Faculty, Gdańsk University of Technology, Poland

Abstract. The chapter presents a method for automatic enhancement of computer character animation utilizing fuzzy inference. First the user designs a prototype version of animation, with keyframes only for important poses, roughly describing the action. Then animation is enriched with new motion phases calculated by the fuzzy inference system using descriptors given by the user. Various degrees of motion fluency and naturalness are possible to achieve. The proposed algorithm of the animation enrichment based on fuzzy description is thoroughly presented. Animation parameterization is presented, new parameters are designed, and the relation between the coefficients proposed and subjective features of motion are established. The first part of the processing consists in creating fuzzy rules for the algorithm based on results of subjective evaluation of the animated movement. The second one utilizes input descriptors for new motion phase calculation, which are finally added to the animation. Quality and fluidity increase of the obtained animation are verified by subjective evaluation. Efficiency of the algorithm proposed is discussed and creation of the animation with the system engineered is compared with animation prepared by utilizing typical computer animation methods.

Keywords: animation, fuzzy processing, subjective evaluation.

1 Introduction

Animated computer characters are often used for virtual reality (VR) applications, computer games, educational software, and serve as actors in animated movies. The animated character portraying a user in the VR is called avatar. Current trends in VR aims at providing full interaction and personalization of avatar's look, outfit, gender, or age [12]. Therefore adding another aspect of personality, i.e. movement style for adjustment is in our opinion a logic step in avatars' development. The most advanced method for acquiring animated movement is Motion Capture, alas it has very high technical requirements [11]. Other drawback of this method is motion representation being unfriendly for animator and hard to edit by hand, in which rendering correction of mistakes is almost unachievable. A very important disadvantage of that method is that capturing motion of a real actor does not allow for achieving exaggerated animation, typical for animated movies and cartoons. Moreover changing of the animation

G.A. Tsihrintzis and L.C. Jain (Eds.): Multimedia Services in Intel. Environ., SIST 3, pp. 103–118.
springerlink.com

style is practically impossible. VR application creators can capture a few variants of action for different types of avatars, but the work would be very tedious, and seems impractical. Therefore new methods should be developed for processing of motion, either already captured by a real actor or a prototype action prepared by the animator, equipped with a new quality and a wide range of motion styles defined by the user.

As seen from the literature sources [2][4][5][7][8][17], animation methods are developed aiming at achieving interesting and high quality motion without motion capture, allowing generation of motion variants e.g. depending on physical parameters in simulation, but no significant progress is currently achieved in this domain.

In the paper a new method for creating animated movement is proposed, combining computer-based with the traditional cartoon animation [1][16]. It is the continuation of research described in the earlier work by the authors [13][14][15]. Resulted from this research work a fluid, high quality animated movement is achieved, with a style spanning between natural and exaggerated depending on the user input. Advanced computer animation techniques employed in the authors' approach guarantee effectiveness in designing an animated movement while utilizing well-known keyframe approach. On the other hand traditional animation approach serves as a source of practical rules for correct utilization of motion phases, assuring fluency of motion, readability of message, and correct display of character personality. That knowledge is available in the animation literature, gathered during last hundred years of animation evolution [1][16].

The new method proposed and verified in our work combines both mentioned domains of animation techniques utilizing at the same time fuzzy processing. The main goal of our work is to create an animation framework serving as a tool for animation preparation and employing fuzzy logic for automatic alteration of its style and fluidity depending on the user's input. Fuzzy logic is here used for modeling the animation knowledge represented as relations between various motion parameters. Needed knowledge is obtained from literature and gathered during data mining of parameters from sample animations. Processing of that knowledge allows for establishing such parameters that have the strongest influence on subjective motion features, i.e. its *quality*, *fluidity* and *style*. Obviously, fuzzy logic enables to operate on non crisp values, imprecise terms, and to process linguistic variables related to subjective descriptors originating in human perception of motion features. Using fuzzy logic in that scope is, according to our knowledge, a unique approach to the computer animation.

2 Animation Parameterization

For further analysis we proposed a model of animation creation process, presented in Fig. 1. The input data for the model are: director/animator requirements related to action, timing, and placing of animated objects (parameters in vector **A**) and requirements related to quality (parameters in vector **Q**). The mechanism of the character movements comprises mainly limbs, therefore the animated objects are character's limbs. Input parameters **A** and **Q** describe sequence of the character's poses defining the action. A working version of the animation is prepared by the animator, fulfilling technical requirements defined by **A**, although it is also accompanied with **additional motion phases** not defined by the director (parameters of additional phases are stored in vector **B**). Additional phases do not change the main action, but add subtle variations to transitions between poses, influencing character *personality*, and motion *fluidity*, *quality*, and *style*. The working version is then subjectively evaluated, and the resulted quality features are stored in vector $\mathbf{Q}^*$ for comparison with the given $\mathbf{Q}$ being expected to be fulfilled by the final animation. If the comparison results are not satisfactory, the animation is corrected, i.e. motion features related to given requirements stored in **A** are not modified, but additional phases proposed by animator stored in **B** are changed. The animator's task is to decide which parameters of **B** should be changed for achieving best results. The process is repeated until the worked-out animation has features close to the assumed ones, then it is sent as the final version at the output. In our method that task is accomplished by fuzzy inference.

The described decision process has some constraints and requires that: (1) motion parameters stored in vector **A** are not changed during the preparation of output animation, i.e. sequence of poses in time and space remain intact, and (2) depending on the assumed action described in vector **A**, parameters stored in **B** should be changed in such a way that the quality requirements stored in **Q** are fulfilled as precisely as possible (see Fig. 1). The subjective evaluation function seen in Fig. 1 (upper/right hand part of Figure) is described in detail in Section 3.1.

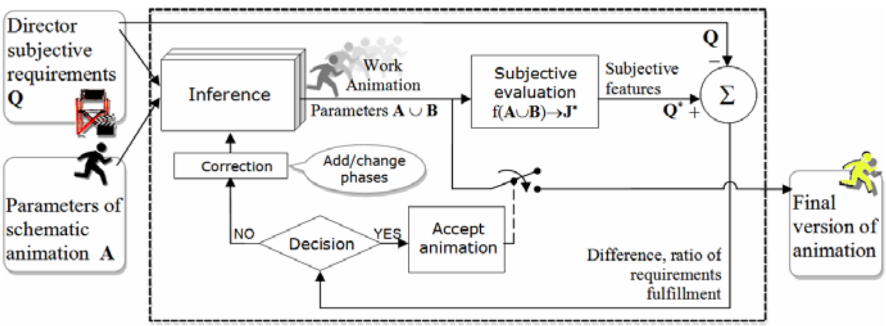


Fig. 1. Animation creation process

Adding new keyframes does not change the objective action of the character, but influences subjective nature of the motion. This fact is used by the animator, as he adds phases of *anticipation* (a motion preceding main transition between two poses adjacent in time), *overshoot* (a motion after transition, when motion stops slowly not abruptly), and *moving hold* (subtle changes of a pose while actions stops, and a character should remain alive). Anticipation displays preparation for the action, e.g. a squat before jump, an overshoot portrays inertia, and moving hold is responsible for showing balancing and maintaining aliveness. Correct utilization of these additional phases influences naturalness and fluency of motion, related to high subjective quality of animation.

For simplification of the problem the following assumptions have to be made:

1. *anticipation* and *overshoot* are alike, i.e. their times (t) and limb rotation amplitudes (A) are assumed to be equal, therefore for parameterization only two values are used: $\mathrm{d}A=A_a=A_o$ and $\mathrm{d}t=t_a=t_o$;
2. times and amplitudes of these phases are limited and cannot extend beyond subjectively accepted values: $\mathrm{d}t\in(0; 10)$ [frames], $\mathrm{d}A\in(0; 0{,}265\pi)$ [rad];

3. *moving holds* are calculated as random movements, with times and amplitudes taken from the limited ranges, however this phase is not in the scope of this work.

The processed motion parameters for a single segment of animation are presented in Fig. 2. Details of the animation segmentation performed at the preprocessing stage of the algorithm are not in the scope of the paper, however they are briefly described at the end of the chapter. Parameters related to poses, taken from the director'/animator's requirements, are stored in vector **A**. Parameters connected to subjective features and related to additional motion phases are stored in vector **B**:

$$\mathbf{A}=[V, A, t] \qquad \mathbf{B}=[\mathrm{d}A, \mathrm{d}t] \tag{1}$$

where: $A=a_3-a_2$, $t=t_3-t_2$, $V=A/t$.

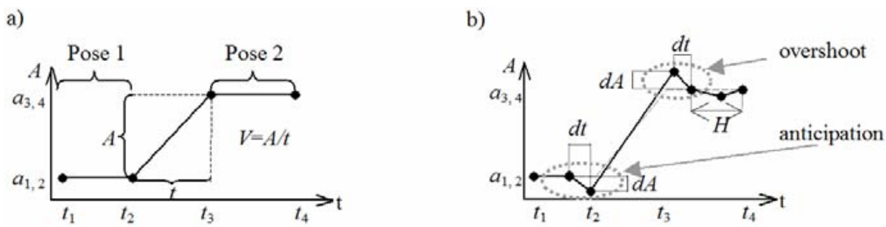


Fig. 2. Motion data: a) transition between two poses is described with parameters **A**=[*V*, *A*, *t*], b) additional phases of anticipation, overshoot and moving hold, 'H' depicts range of subtle random variation of the target pose.

3 Motion Data Processing

Traditional animation rules [1][16] describe a way of the additional phase utilization for achieving particular subjective results. Namely, fast motion with big amplitude should be preceded by the large anticipation and finished with the large overshoot; long motion should be preceded with a long anticipation and finished with a long overshoot. Taking this into consideration it was assumed that the proportionality occurs between these parameters, which can be described as (2):

$$\mathrm{d}A=alpha\cdot V\cdot A \qquad \mathrm{d}t=beta\cdot V\cdot t \tag{2}$$

where *alpha* and *beta* are proposed proportionality coefficients, which will be discussed later. In case when d*A* or d*t* exceeds the assumed maximal value, following *f*(*x*) and *g*(*x*) saturation functions are used (3):

$$\mathrm{d}A=f(alpha\cdot V\cdot A),\ \text{where}\ f(x)=0{,}265\pi\cdot\mathrm{tgh}(x / 0{,}22\pi) \tag{3}$$
$$\mathrm{d}t=g(beta\cdot V\cdot t),\ \text{where}\ g(x)=10\cdot\mathrm{tgh}(x / 0{,}125)$$

It was verified during visual tests what maximum values of amplitudes d*A* and times d*t* are subjectively accepted by the viewers.

Considering Eqs. (2) and (3) variables dA and dt depend only on coefficients *alpha* and *beta*. If dA and dt changes have an influence on subjective meaning of the animation, then a relation should also exist between subjective features *alpha* and *beta*. Relations between objective numerical values of *alpha*, *beta* and subjectively evaluated animation features: *style* and *fluidity* were specified in data mining of results coming from subjective test. In evaluation tests a simple animations where used, containing two poses and a transition, with additional anticipation, overshoot and moving hold phases. Test participants' task was to name features of motion utilizing the following discrete scales: *style*={*natural*, *middle*, *exaggerated*}; *fluidity*={*fluid*, *middle*, *abrupt*}; *quality*= {1, 2, 3, 4, 5}. Results of *fluidity* and *style* evaluation are stored in vector **Q**=[*style*, *fluidity*], and *quality scores* are processed individually as an additional criterion **QS**. Evaluation of visual stimuli was performed with respect to recommendations indicated in literature sources [9][10].

First the correlation between *alpha* and *beta* of each evaluated animation and its *style* and *fluidity* scores was calculated. The results are presented in Table 1. Strong correlation indicates that a certain connection between the selected subjective feature and the proposed coefficient exists, therefore rules describing that relation can be created.

The graphical representation of correlations between answers in subjective tests and objective parameters of animations (*alpha* and *beta* coefficients) is presented in Fig. 3.

Table 1. Correlation coefficients between subjective and objective parameters of animations

	beta-style	beta-fluidity	beta-quality	alpha-style	alpha-fluidity	alpha-quality	style-fluidity	style-quality	fluidity-quality
R	-0.14	0.86	0.81	0.82	0.16	0.09	-0.21	-0.27	0.94

3.1 Relations between Animation Features

The correlation described above is used for creating rules that connect objective and subjective parameters of animation. **Ambiguous** information about that relations is obtained during subjective evaluation tests, when participants for an animation described with **A** and **B** select values for subjective features **Q**. A particular animation described by **A**=[V, A, t] and **B**=[dA, dt] may not be evaluated identically by every test participant. Projection from **B** to **Q** with given **A** is denoted as $f_{\mathbf{A}}$: **B**→**Q**. That ambiguous function is called **evaluation function** and reflects viewers' answers. It is assumed that **A** and **Q** are constant and only **B** parameters are unknown variables. Therefore inverse function is being sought, $f_{\mathbf{A}}^{-1}$: **Q**→**B**, which for given **A** and required values of features **Q** picks correct **B**. That function is also ambiguous. For each **A** and **Q** a result is first generated as a set of objects – animations that have an action as the one defined in **A** and that were subjectively evaluated as having values matching given **Q**=[*style*, *fluidity*], but are differentiated by **B**. From this set one object is finally selected based on additional criterion – maximization of *mean quality score* **QS**. Therefore for any **A** it is possible to generate **unambiguous** rules connecting values of given subjective requirements **Q** with parameters of additional phases **B**. These rules describe what values of additional phases should be used if the working

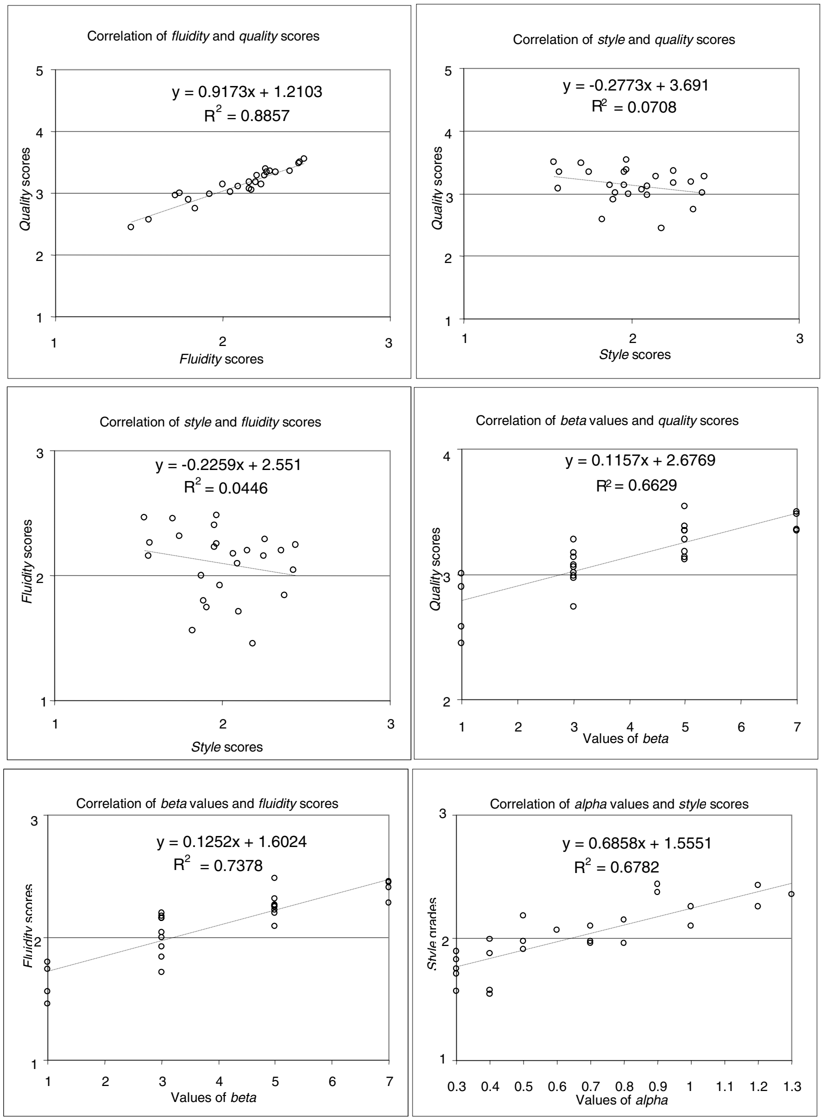


Fig. 3. Relations between subjective scores and proportionality coefficients *alpha* and *beta*

animation is described by values stored in vector **A** and animation subjective features values should match those given in **Q**.

Fig. 4 represents the graphical representation of the unambiguous result searching problem described above. Discrete hyper cubes of parameters visualize finite and discrete domains of multidimensional parameters vectors **Q** and **B**. A given **Q** value for the inverse function $f_{\mathbf{A}}^{-1}$: $\mathbf{Q} \rightarrow \mathbf{B}$ points to various **B** values and the result is ambiguous. On the other hand for any **B** there exists an additional mediation value **QS**, quality

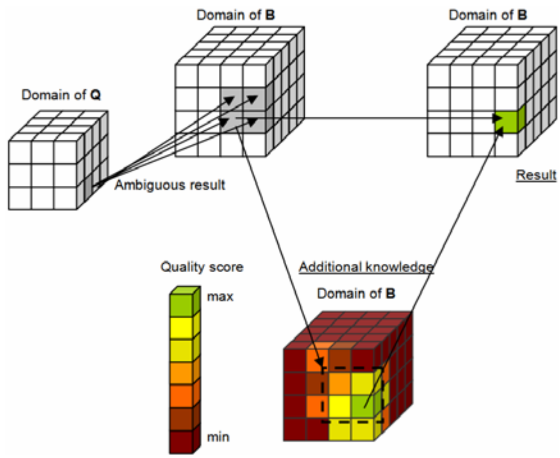


Fig. 4. Illustration of unambiguous result searching

score obtained in the subjective evaluation test. Maximum **QS** is sought for all **B**s obtained in first step (area marked by dashed line), and only one result **B** is selected.

Equations (2) and (3) describe proportionality between values V, A and t stored in vector **A**, and parameters of additional motion phases dA and dt stored in vector **B**. Coefficients of these proportionalities, i.e. *alpha* and *beta*, are used for simplification of searching for $f_{\mathbf{A}}^{-1}$: $\mathbf{Q} \rightarrow \mathbf{B}$. The problem is first reduced to defining relations $\mathbf{Q} \rightarrow$ [*alpha*, *beta*], then based on (3) inverse function $f_{\mathbf{A}}^{-1}$: $\mathbf{Q} \rightarrow \mathbf{B}$ is calculated for any given **A**. Obtained relations between subjective variables and proportionality coefficients are presented in Table 2.

Table 2. Calculation of *alpha* and *beta* based on given *fluidity* and *style*

alpha		*fluidity*		
		abrupt	middle	fluid
style	natural	0.7	0.5	0.3
	middle	0.9	0.7	0.5
	exaggerated	1.3	1.1	0.9

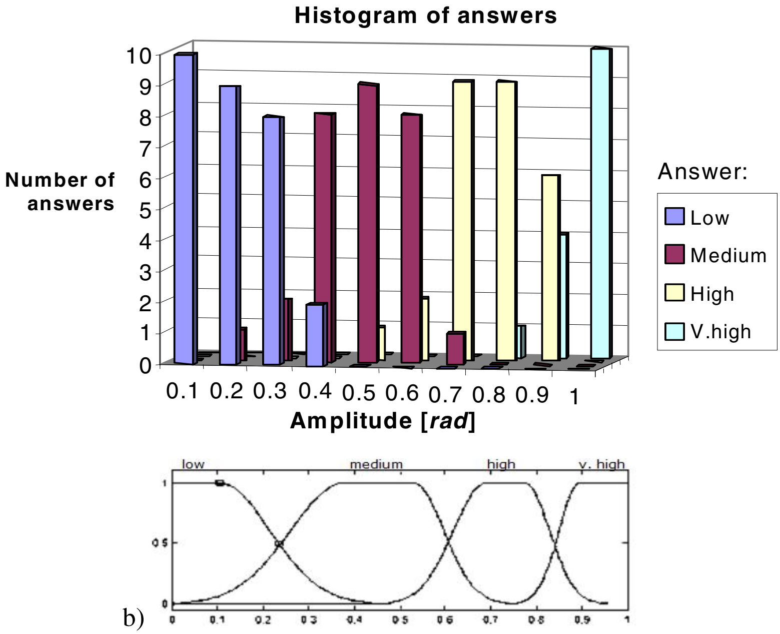
beta		*fluidity*		
		abrupt	middle	fluid
style	natural	3	5	7
	middle	1	5	5
	exaggerated	3	5	7

3.2 Generation of Rules

Based on the knowledge gathered during performed subjective tests a set of rules was obtained, and implemented in fuzzy inference system which replaces analytic Equations (3):

$$\begin{array}{l} \text{IF } V{=}\ldots\wedge t{=}\ldots\wedge style{=}\ldots\wedge fluidity{=}\ldots \text{ THEN d}t{=}\ldots \\ \text{IF } V{=}\ldots\wedge A{=}\ldots\wedge style{=}\ldots\wedge fluidity{=}\ldots \text{ THEN d}A{=}\ldots \end{array} \quad (4)$$

For all parameters *fuzzy membership functions* are required, therefore input parameters V, A, t are first discretized: V={0.0, 0.05, 0.1, 0.15,...,0.4}, A={0.1, 0.2,...,1.0}, t={5,10,15,…,50}. Then calibration animations are prepared for evaluation, presenting an animated character arm motion with speed, amplitude or time chosen as one of the above discrete values. Features of these animations are rated utilizing linguistic descriptors: *speed*={*low, medium, high, very high*}, *amplitude*={*low, medium, high, very high*}, *time*={*short, medium, long, very long*}. Based on evaluation results membership functions (mf) are created. For example for mf *speed*=*low* as a kernel[1] a range of discrete values of feature *speed* for which participants selected linguistic value *low* more often than in 80% of cases is selected (Fig. 5a.). Membership functions for variable amplitude are presented in Fig. 5b.



a)

b)

Fig. 5. Relation between number of answers and membership functions for variable *amplitude*: a) histogram of scores for amplitude values of basic motion, b) membership functions created on the basis of the number of answers

Crisp values obtained by calculation of Eq. (3) for all combinations of discrete values of input parameters *V*, *A*, *t* were also fuzzified. In the first step all crisp values were clustered using *k*-means algorithm, then triangle mf were created, each having maximum in the center value of the respective *k*-th cluster, and spanning in a way that will fulfill sum-to-one condition, as is presented in Fig. 6.

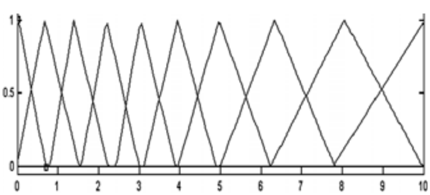


Fig. 6. Obtained membership functions for variable d*t*

[1] Values with membership equal 1.

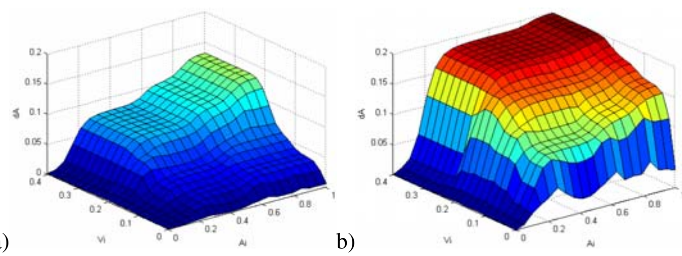
Fuzzy rules (4) are formulated for all three linguistic values of variables *style* and *fluidity* and for all four values of discretized V, A, t. First, based on Table 2, for the given *style* and *fluidity* coefficients *alpha* and *beta* are calculated, then from Eq. (3) the outcome values are calculated which are finally fuzzified. Repeating that process for all combinations of input values all rules required are being formulated. Examples of rules for calculating fuzzy dA values are presented in Table 3.

Table 3. Rules for calculating fuzzy value of dA for the given *natural style* and *fluid animation*

dA		V			
		low	medium	high	v. high
A	low	mf1	mf1	mf2	mf3
	medium	mf2	mf2	mf3	mf5
	high	mf2	mf3	mf4	mf6
	v. high	mf2	mf3	mf4	mf7

mf$n$ – n-th membership function for variable dA

The analysis of interpolation surfaces of fuzzy system (Fig. 7) reveals that the system is complete, i.e. a result exists for any combination of input parameters, and changes of the output value are fluid (continuous first derivative), which are important features for practical usage of the system.



a) b)

Fig. 7. Interpolation surfaces of fuzzy system – dA values depending on input V and A for: a) natural style and fluid animation, b) exaggerated style and fluid animation.

The processing described above is performed on a single segment of the animated movement, i.e. starting with held pose, then transitional movement(s) and finishing with held pose, as was presented in Fig. 2. For processing the animation should be correctly segmented first. Fig. 8 shows animation with 8 poses where only 1st, 2nd and 8th pose have a hold phase and therefore depicts the start and the end of a segment. Single additional anticipation and overshoot phases are inserted respectively at the beginning and at the end of a segment, based on parameters **A** of the transitional movement and the given descriptive values of *style* and *fluidity*. In that way for any animation containing distinctive poses the user can define required subjective features of the motion and after the processing get the enhanced animation with high quality and a personalized movement.

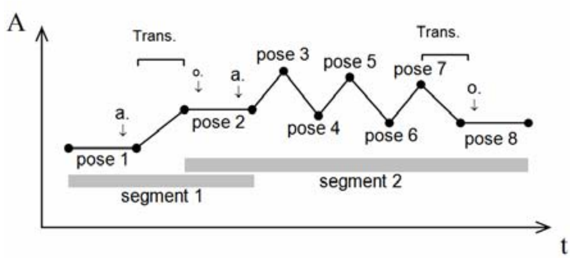


Fig. 8. Segmentation of animation. Arrows denote places for insertion of anticipation (a.) and overshoot (o.), transitional movements started or finished with held pose are marked. These transitional movements are parameterized with values stored in vectors **A**.

4 Results

The effectiveness of the animation processing system was verified in visual subjective tests. Five prototype animated avatar actions served as a test material. These were enhanced using the system and all combinations of input descriptors for *style* and *fluidity*. All animations were rated using 5-point scale. For verification also non-processed versions of animations were displayed. Mean Opinion Score values for all animations are presented in Table 4.

Table 4. Mean opinion score for animations depending on values of subjective descriptors used for enhancement

Mean opinion score		*style*		
		natural	medium	exaggerated
fluidity	abrupt	2.08	2.06	2.08
	medium	3.28	3.22	3.1
	fluid	4.24	3.92	4.02
Non-processed			1.5	

The processed animations obtained statistically valid higher scores than non-processed (Fig. 9). Moreover the fluid motion was always rated higher than the abrupt

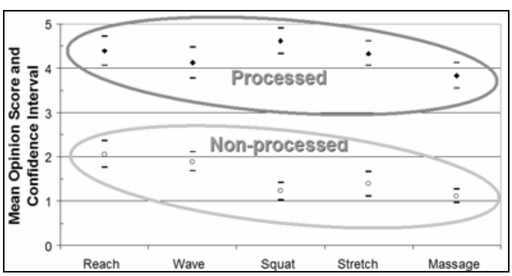


Fig. 9. Mean Opinion Score values with confidence intervals for processed and non-processed animations of five actions utilizing values *style*: *natural*, *fluidity*: *fluid*

one (Fig. 10). Finally, variation of animation style does not have an influence on quality scores (Fig. 11), therefore the developed method for animation creation can by applied for generating many versions of a single prototype action, matching user's requirements for *style* and *fluidity*.

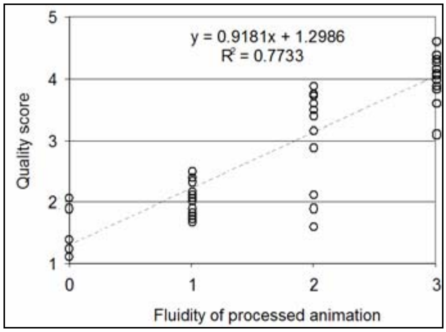


Fig. 10. Correlation between fluidity of animation and obtained scores for all five test animations. Fluidity "0" - denotes *non-processed*, "1" - *abrupt*, "2" - *medium*, and "3" - *fluid*.

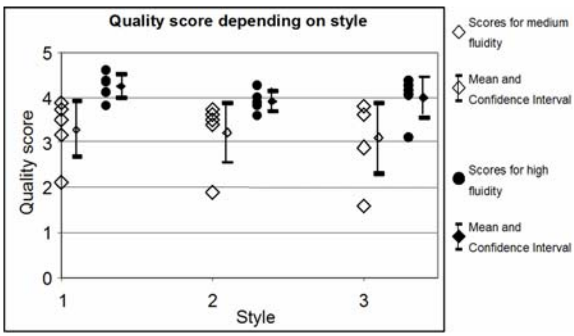


Fig. 11. Quality scores depending on style for two levels of fluidity: medium and high. Particular scores and mean scores are shown. Style "1" denotes *natural*, "2" - *medium*, and "3" - *exaggerated*.

4.1 System Efficiency

The efficiency of the animation creation process utilizing the designed system is higher comparing to utilization of typical computer animation software only. In our approach the animator is expected to design keyframes only for each pose in the action sequence. Other keyframes i.e. anticipations, overshoots and moving holds are inserted by the system automatically. More than 50% of all keyframes are the additional ones, therefore at least half of the animator's work is done automatically. Let us assume that the animation with 10 poses and short moving holds, where only one oscillation keyframe will be inserted, contains:

- b·(10 keyframes for poses starts + 10 keyframes for poses ends) = b·20
- b·[(10–1) phases of transitions between poses. For each there is one anticipation keyframe, and one overshoot keyframe] = b·18
- b·[(10–1) moving holds after each transitions. For each there is one keyframe of oscillation] = b·9,

where b is number of character limbs (*bones*).

Total sum of keyframes is b·20 for the animator and b·27 for the system, added automatically.

If animation contains P poses and N frames of oscillation for each moving hold, then the ratio of the number of keyframes created by the animator to the number of keyframes added automatically by the system can by formulated in the following way:

$$\frac{animator\ keyframes}{system\ keyframes} = \frac{2P}{(2+N)(P-1)} \tag{5}$$

If the animation is long, then P is large, and (5) can be simplified to:

$$\frac{animator's\ keyframes}{system\ keyframes} = \frac{2P}{(2+N)(P-1)} \approx \frac{2P}{(2+N)(P)} = \frac{2}{(2+N)} = 1 + \frac{2}{N} \tag{6}$$

Creation of P poses based on exact instructions from the director and then utilizing the proposed system is much simpler and a less demanding task than a long and cyclical introduction of new keyframes, with various time and amplitude values, evaluation of subjective meaning of the motion, and repetition of the process for achieving correct artistic effect.

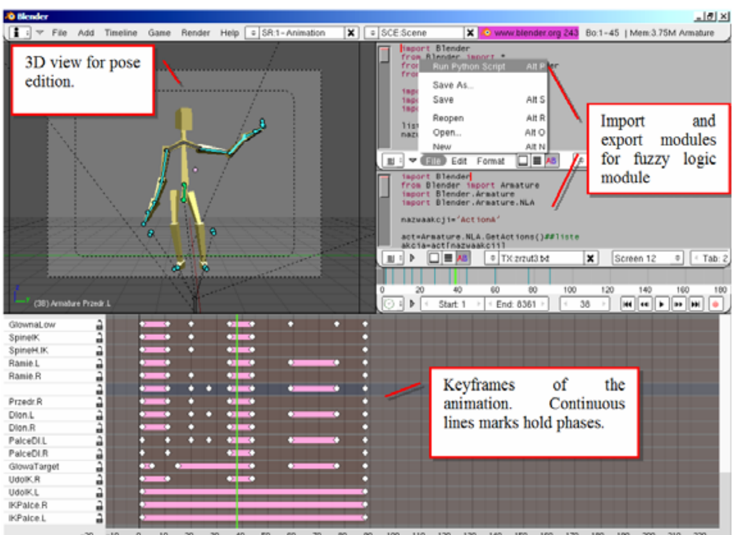


Fig. 12. User interface of animation creation system

4.2 Animation System

The technical implementation of the animation enhancement system involved 3D graphics and animation software Blender [6], for which a "bridge" application was prepared, connecting animation interface with the fuzzy logic inference system in MATLAB. During utilization of the system the workflow is as follows. An animation is designed in a working version, containing poses and fulfilling director's requirements, then animation data are being sent to the inference module. That module provides the user interface for defining values of required *fluidity* and *style* of animation. Finally, when a fuzzy processing is performed, new animation data are transferred back into 3D animation software, where the animation is visualized. The user interface for the animation creation is presented in Fig. 12. The system is designed in such a way that it provides the possibility of the rule set extension and introducing new descriptive parameters of motion in the future.

5 Conclusions

In this Chapter the system for animation enhancement was presented based on methodology and rules coming from traditional animation combined with fuzzy processing. In the system fuzzy rules are used for calculation of parameters of additional motion phases that should be inserted to animation for enhancement of its subjective quality and to introduce new subjective features such as stylization and fluidity of motion. New proportionality coefficients *alpha* and *beta* were defined that are strongly correlated with subjective features of animation.

Results of animation subjective evaluation tests show that processed animations are rated much higher than non-processed for their quality and motion fluidity. Moreover processing of simple animation prepared by nonprofessional animator, always leads to significant increase of its quality. Comparing processes of the animation creation with and without the system the utilization of the proposed method is always more time-effective.

Original results include:

- Defining proportionality coefficients *alpha* and *beta* that are strongly correlated with subjective features of animation.
- Defining fuzzy membership functions for animation parameters corresponding with viewers' subjective motion perception.
- Developing a system for the animation enhancement, utilizing fuzzy inference in a novel way for computer animation. The effectiveness of the system was verified subjectively by a large group of viewers.

The developed method for animation enhancement can be applied to generation of many versions of a single base motion, matching user requirements for style and fluidity, useful in VR applications where avatar's behavior personalization is important. It can be applied also for fast creation of high quality animated movement, when the user has only basic knowledge of animation and prepares only main poses, and additional motion phases are added automatically by the system. Software for automatic creation of animation may be envisioned, where kind of meta language can be used to define hierarchically animation as a set of actions, action as a set of poses, and target fluidity and style as additional parameters. Main poses can be then selected from the database of poses and fluidity and style modifications applied by the system. That kind of automatic animation can be utilized in the educational software, interactive virtual instructors or helpers, communication software, entertainment software, and computer games.

Resources

A Key Books

- Blair P (1995) *Cartoon Animation*. Walter Foster Publishing, Laguna Hills
- Magnenat-Thalmann N, Thalmann D (1996) Interactive Computer Animation, Prentice Hall
- Menache A (1999) *Understanding Motion Capture for Computer Animation*. Morgan Kaufmann, San Francisco
- Mullen T, Roosendaal T, Kurdali B (2007) *Introducing Character Animation with Blender*. Sybex, USA. See also web page: http://www.blender.org
- O'Rourke M (1995) *Principles of Three-Dimensional Computer Animation*, W. W. Norton and Company, New York,
- Parent R (2001) *Computer Animation: Algorithms and Techniques*, Morgan-Kaufmann, San Francisco. On-line version: *http://www.siggraph.org/education/materials/HyperGraph/animation/rick_parent/Outline.html*
- Pocock L, Rosebush J (2001) *The Computer Animator's Technical Handbook*, Morgan Kaufmann, San Francisco
- Thomas F, Johnston O (1981) *Disney Animation – The Illusion of Life*. Abbeville Press, New York
- Vince J (1992) *3-D Computer Animation*, Addison-Wesley, New York
- Watt A, Watt M (1992) *Advanced Animation and Rendering Techniques*, Addison-Wesley, New York

B Key Survey/Review Articles

- Lasseter J, *Principles of Traditional Animation Applied to 3D Computer Animation*, Computer Graphics, pp. 35-44, 21:4, July 1987 (SIGGRAPH 87). On-line version: *http://www.siggraph.org/education/materials/HyperGraph/animation/character_animation/principles/prin_trad_anim.htm*
- Vercher JL, *Perception and Synthesis of Biologically Plausible Motion: From Human Physiology to Virtual Reality*. Lecture Notes in Artificial Intelligence LNAI 3881, Springer-Verlag Berlin Heidelberg, 2006

C Key Journals

- *Computer Animation and Virtual Worlds*, Wiley
- *Journal of Visualization and Computer Animation*, Wiley

D Key International Conferences/Workshops

- ACM SIGGRAPH, *http://www.siggraph.org/s2009/*
- Computer Graphics, Visualization, Computer Vision and Image Processing, Portugal, *http://www.cgv-conf.org/*
- GRAPP, International Conference on Computer Graphics Theory and Applications, *http://grapp.visigrapp.org*

E Software

- Blender 3D, free, open source, computer animation and graphics software: *http://www.blender.org/*
- MakeHuman, human modelling software based on Blender 3D: *http://www.makehuman.org*

F Web Pages

- Blender news and tutorials: *http://www.blendernation.com/*
- BlenderART magazine, sensitive information and tutorials: *http://www.blenderart.org/*
- Computer animation SIGGRAPH educational materials: *http://www.siggraph.org/education/materials/HyperGraph/animation/anim0.htm*
- Humanoid animation standard (ISO/IEC FCD 19774): *http://h-anim.org/*
- Computer animation resources: *http://www.cse.ohio-state.edu/~parent/animation/index1.html*
- Current trends in computer graphics and animation: *http://www.cs.wisc.edu/graphics/Gallery*

References

[1] Blair, P.: Cartoon Animation. Walter Foster Publishing, Laguna Hills (1995)

[2] Dontcheva, M., Yngve, G., Popović, Z.: Layered Acting For Character Animation. In: Proc. SIGGRAPH, Los Angeles, pp. 409–416 (2003)

[3] Lasseter, J.: Principles of Traditional Animation Applied to 3D Computer Animation. In: Proc. SIGGRAPH, Los Angeles, pp. 35–44 (1987)

[4] Li, Y., Gleicher, M., Xu, Y., Shum, H.: Stylizing Motion with Drawings. In: Proc. SIGGRAPH, Los Angeles, pp. 309–319 (2003)

[5] Milenkovic, V.J., Schmidl, H.: Optimization-Based Animation. In: Proc. SIGGRAPH, Los Angeles, pp. 37–46 (2001)

[6] Mullen, T., Roosendaal, T., Kurdali, B.: Introducing Character Animation with Blender. Sybex, USA (2007), `http://www.blender.org`

[7] Neff, M., Fiume, E.: Artistically based computer generation of expressive motion. In: Proc. Symposium on Language, Speech and Gesture for Expressive Characters, pp. 29–39 (2004)

[8] Popović, J., Seitz, S.M., Erdmann, M.: Motion Sketching for Control of Rigid-Body Simulations. Association for Computing Machinery Transactions on Graphics 22(4), 1034–1054 (2003)

[9] RECOMMENDATION ITU-R BT.500-11, Methodology for the Subjective Assessment of the Quality of Television Picture (2002)

[10] RECOMMENDATION ITU-T P.800, Methods for Subjective Determination of Transmission Quality. Geneva, Switzerland (1996)
[11] Remondino, F., Schrotter, G., Roditakis, A., D'Apuzzo, N.: Markerless Motion Capture from Single or Multi-Camera Video Sequence. In: Int. Workshop on Modelling and Motion Capture Techniques for Virtual Environments (CAPTECH), Zermatt, Switzerland (2004)
[12] Weber, A., Rufer-Bach, K., Platel, R.: Creating Your World: The Official Guide to Advanced Content Creation for Second Life, Sybex, USA (2007), `http://www.secondlife.com/`, `http://wiki.secondlife.com/wiki/Animation`
[13] Szczuko, P., Kostek, B.: Analysis and Generation of Emotionally Featured Animated Motion. In: Ślęzak, D., Yao, J., Peters, J.F., Ziarko, W.P., Hu, X. (eds.) RSFDGrC 2005. LNCS (LNAI), vol. 3642, pp. 333–341. Springer, Heidelberg (2005)
[14] Szczuko, P.: Application of Fuzzy Rules in Computer Animation. PhD Thesis. Gdańsk University of Technology, Poland (2008)
[15] Szczuko, P., Kostek, B.: Enhancement of Computer Character Animation Utilizing Fuzzy Rules (in reviews). In: 2nd International Symposium on Intelligent Interactive Multimedia Systems and Services, KES IIMSS 2009, Mogliano Veneto (2009)
[16] Thomas, F., Johnston, O.: Disney Animation – The Illusion of Life. Abbeville Press, New York (1981)
[17] Vercher, J.L.: Perception and Synthesis of Biologically Plausible Motion: From Human Physiology to Virtual Reality. In: Gibet, S., Courty, N., Kamp, J.-F. (eds.) GW 2005. LNCS (LNAI), vol. 3881, pp. 1–12. Springer, Heidelberg (2006)

Chapter 6

Instructional Design of a Requirements Engineering Education Course for Professional Engineers

Takako Nakatani[1], Toshihiko Tsumaki[2], and Tetsuo Tamai[3]

[1] University of Tsukuba, Graduate School of Business Sciences
nakatani@gssm.otsuka.tsukuba.ac.jp
[2] National Institute of Informatics
tatsumaki@nii.ac.jp
[3] The University of Tokyo
tamai@graco.c.u-tokyo.ac.jp

Abstract. As offshore development has become common for software companies, those companies have started to concentrate their engineers' effort upon the early development phase of software. Within this context, it is important to educate professional engineers to master all aspects of requirements engineering. It is not easy to master the various aspect all at once. However, we can expect professional engineers to study methods on their own if they believe that learning the methods is important for them or their projects. We have developed a two-and-a-half-day role-playing workshop for professional engineers that focuses on teaching the importance of requirements engineering, as well as the background, rationale, and purpose of the requirements to guarantee success in their projects. We start this paper with an overview of requirements engineering and its techniques, and then, introduce the instructional design of our course for professional engineers. We also present the results of an actual workshop, which showed engineers could earned clients' mindsets and the importance of the rationale of requirements.

Keywords: Requirements Engineering, Engineer Education.

1 Introduction

As offshore development has become common for software companies, those companies have started to concentrate their engineers' effort upon the early development phase of software.

Nissan Techno Vietnam, which Nissan Techno set up in 2001 to address the shortage of talented engineers in Japan, is just one example of offshore development. The cooperative relationships between Nissan, Nissan Techno, and Nissan Techno Vietnam are shown in Figure 1. Each company shares the mission of supplying high-quality cars, but each plays different role. For example, Nissan concentrates on "what" to research and develop for their next technology, while Nissan Techno Vietnam implements in their various cars the technologies standardized by Nissan Techno Japan.

The software industry is also facing the shortage of talented engineers. If we draw a comparison between the Nissan example and the software industry, the research and

G.A. Tsihrintzis and L.C. Jain (Eds.): Multimedia Services in Intel. Environ., SIST 3, pp. 119–151.
springerlink.com

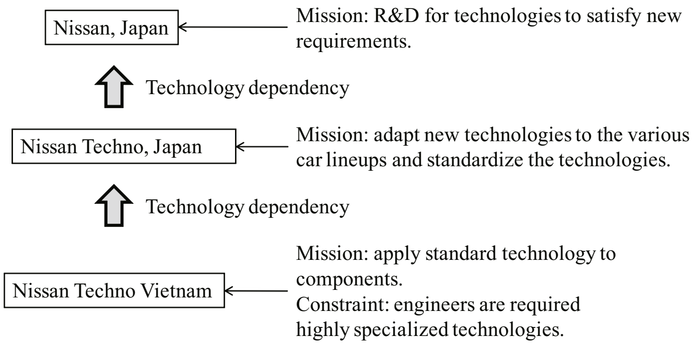


Fig. 1. Technical responsibilities and organizations

development activity of the automobile industry can be equated to the early requirements phase that focuses on the "what" aspect of software development [1,2,3]. Indeed, if software vendors can procure offshore developers, they can shift in-house engineers from the middle and/or later development phases to the early development phase.

In this current situation, one problem is that engineers, including managers, still focus on the middle to later development phases. The importance of requirements engineering (RE) has been recognized since Boehm's paper [4]. His work showed that correcting requirements errors on large projects is 100 times more expensive in the requirements phase than in the requirements phase. However, most engineers do not prioritize RE because they believe there is not enough time to practice RE while trying to finish their jobs on time. Here is an interesting anecdote from a project manager about RE.

> There is an engineer who can produce thick software requirements specifications files from a plentiful amount of information concerning stakeholders' real requirements, but he can also produce the same volume of specifications from a small set of stakeholders' requirements.

This means that engineers can form software requirements specifications (SRSs) independently of the stakeholders' needs. What makes this aspect of engineering culture even worse is that no reviewer can distinguish the latter from the former. If the rationale for each requirement is described in the SRS, the reviewers may be able to distinguish one from the other. However, most SRSs contain little description of the background and/or purpose of the requirements. Under these circumstances, unsuccessful elicitation of requirements is the main cause for project delays and this is a difficult problem to overcome. How can we cope with this situation? How should we educate engineers to interest them in RE and motivate them to write proper SRSs? We must develop educational courses to change their culture.

We chose engineers with over ten years of experience as our target students to simplify the problem. Figure 2 charts the typical career paths of Japanese software engineers. After approximately ten years of development experience they become analysts. This paper refers to such analysts as professional engineers. Consequently, we can assume that professional engineers have programming, design, and communication skills, as well as team-work skills. This assumption is unlikely to be applicable to undergraduate

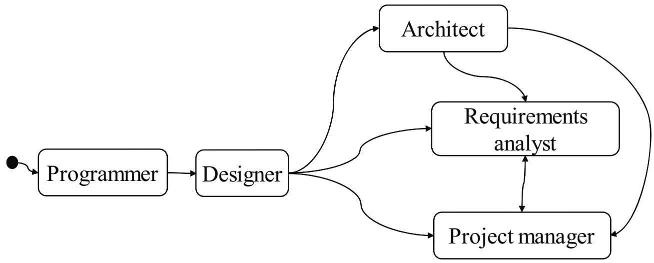


Fig. 2. Typical career paths of Japanese engineers

students, although RE learning at the undergraduate and graduate levels is, of course, important [5]. Teaching those skills to students remains a challenge for us. However, that is outside the scope of this paper.

Here is another cultural problem. Professional engineers tend to be suspicious of new techniques. They will make an effort to master a new technique if, and only if, it looks useful to them. The purpose of our educational course is to motivate them to apply RE techniques to their job. Therefore, we had to focus on how to motivate them. This paper introduces an educational course to address these problems.

The remainder of this paper is organized as follows. Section 2 introduces some common RE educational styles. Section 3 introduces the materials used for our RE education course. It includes the definition of requirements engineering and the quality characteristics of SRSs, as well as a project's environment to which professional engineers contribute. Most software projects receive a lot of requirements changes after the external design phase. We present the requirements elicitation process of a real development project. This is a typical situation that professional engineers have to cope with. Then, we introduce an RE technology map classifying the RE techniques. Section 4 defines the instructional design of the course. We discuss the purpose of the course, a metamodel for the course products, some selected methods, and the course style, as well as its process. Section 5 introduces our course experience. A course evaluation is also presented in this section. The final section presents our conclusions.

2 Educational Styles for RE

A lot of RE education courses have been developed. Each style of classroom education and workplace education is useful in its own way [6].

The basic style is a lecture-based course. This style is suitable for an introduction to RE. The basic RE method, BaRE [7], was designed to cover all the key issues in RE in order to introduce the knowledge necessary to use RE.

There is another style of educational course developed by Lami [8]. It was developed for undergraduate students to provide them with the specific abilities and skills to face the demands of the professional software engineering market. This 16-hour course is composed of several modules: software quality, requirements documentation, a requirements elicitation workshop, and quality analysis techniques, methods, and tools.

Berenbach et al. have introduced a unified requirements model that integrates features, use cases, and hazard analysis [9]. It is important that engineers learn multiple methods, not just a single method. Furthermore, the relationships between the methods should be taught so that engineers gain a better understanding of the unified requirements model and how it could replace their current methods.

Most RE education courses regard the students' participation as important. Therefore, game-based, role-playing, experiment-based, investigation-based, and/or workshop styles of teaching have produced good results.

RE-O-Poly [10] is a *Monopoly*-like game designed to teach RE methods and interviews. The players face various situations that RE engineers are regularly confronted with. By playing the game, students can learn ways and places to apply RE methods.

Beatty and Alexander [11] reported their initial experience with game-based RE training. Their games were developed for learning a categorizing technique, requirements models, and the traceability matrix. For example, "silent affinity diagrams" is a silent game played in pairs to teach logical groupings of requirements. Keeping silent makes the students concentrate on the game. Betty and Alexander concluded that games seem to be transferable and applicable to training in RE. Game-based training may be interesting for students. However, we assume that it would not be as successful for the further education of professional engineers. Furthermore, our sponsor company expects us to present a serious course, not a pleasure course.

Students should understand what their stakeholders' want/their stakeholders' needs. To achieve this, a role-playing exercise can be an appropriate style for learning RE [12]. One of the strong points of role-playing exercises is that students can experience different organizational contexts. As a result, they can gain first-hand knowledge of the mindsets of people in roles they would not ordinarily experience. Such an experience is important for professional engineers, because their real roles are always as the developer and they have few chances to appropriately consider their clients' mindsets in their daily lives. Therefore, we decided to apply role-playing exercises to help the students better understand their clients' mindsets.

Regev et al. developed a course for university students [13]. Their educational style is an experiment-based style that uses a low-tech demonstration of a realistic development organization. Through the educational experience, students learn social and/or design-problem complexities. The problems with teaching RE to university students are also presented. Some of them are also common for professional engineers, since they tend to expect an RE environment in which the stakeholders can clearly express their requirements and/or have sufficient domain knowledge. In a real project, of course, engineers are the ones who are responsible for adequately eliciting and defining requirements while acquiring sufficient domain knowledge.

The investigation-based approach aims at adding value to text-based learning [14]. This course was developed for undergraduate students as a one-semester course. The teacher instructs students on the established concepts in RE and then has the students carry out their normal exercises. After that, the students analyze their own data and interpret their findings. The students are expected to discover new knowledge through their investigations.

There are several RE courses developed for professional engineers. Siemens has several RE courses for engineers worldwide [15]. The first one, which teaches the foundations of RE, is conducted as a three-day lecture-based educational course. After the course, the students are required to answer questions concerning the frequency and usage of the information that they learned during the RE course. The most frequently-used technique is reported to be requirements writing. Therefore, our course includes requirements writing.

Simmons [16] recommended that further education through corporate training should include a lecture on the client company's culture and specific requirements, e.g., safety requirements, hardware constraints, and/or tools environment. Simmons also pointed out that the most common problem for professional engineers is the duration limitation. The course was usually less than a week, and only two or three days on average. Therefore, follow-on mentoring was required to strengthen the adoption of new practices and to help ensure first-use success. Our course is followed by a lecture on the company's specific constraints, tools environment, and culture. The students have to make a presentation about their results to their client team and conduct an evaluation of the effectiveness several months after the course.

The viewpoint from which RE education is taught differentiates the various available courses. Engineers areexpected to be involved with business-related issues. Thus, RE education cannot bypass the business aspect of the process. Wegmann developed a course that views IT as a way to enhance the strength and compensate for the weaknesses of a business [17]. It is a problem-based learning course for university students. The course length is 84 periods over one semester. The final goal of his course is for students to discover the importance of managing the trade-offs between technical and business issues. Since the course is conducted for university students, its curriculum includes business thinking, e.g., collaborative work, and sustaining business competitiveness through the use of a competitive game. Furthermore, role-playing to practice conducting an IT requirements analysis is incorporated as an IT integration project. If the students are professional engineers, although it may be easy to teach them business thinking, it is not easy to get them to focus on business goals rather than technical ones.

In our case, most students are directed by their bosses to attend the course as part of the company's strategy for shifting their engineers from the middle development phase to the early development phase. Therefore, the goal with our students is to get them to start learning RE as soon as possible and to apply their skills to their own work. They need motivation to start learning, so we selected an experience-based workshop style. In the workshop, each engineer plays the roles of both requester and analyst in order to understand clients' mindsets as well as to evaluate their techniques from the clients' viewpoints. As the target students of our course are professional engineers, the duration of our course is limited to two-and-a-half days.

3 Teaching Materials

In this section, we introduce several topics covered in RE education.

- Introduction to RE
- Quality of SRS

– Actual project situation for requirements changes
– Means of achieving RE goals

3.1 Introduction to RE

First of all, we need to answer the question of what professional engineers need to learn. Shaw [18] refers to engineering as "the creation of cost-effective solutions to practical problems by applying scientific knowledge to building things." According to this definition, Nuseibeh defined RE as follows [19].

– RE is an engineering process to anchor development activities to real-world problems.
– The purpose of RE is to analyze the appropriateness and cost-effectiveness of the solution.
– It presents a series of engineering decisions that are based on the recognition of a problem to be solved and leads to a detailed specification of that problem.

In RE activities, engineers receive various requirements from the real world and then analyze them to develop an SRS. The SRS specifies a particular software product, program, or set of programs that perform certain functions in a specific environment [20]. The requirements specification is like food in that it provides all the nutrition for the development processes that follow. Insufficient food can only produce an inefficient product. This basic idea of RE can be taught by lecture.

The RE process is composed of the four main phases shown in Figure 3 [21]. The first phase of the RE process is requirements elicitation. Here, a requirements engineer identifies the stakeholders and interviews them to extract and analyze their real world problems. After that, the clients, end users, project sponsors, and developers negotiate to reach a consensus about the problems to be solved. In the second phase, the engineer starts to model the requirements and develops the requirements specifications. The third phase focuses on requirements validation. It includes a feasibility study and impact analysis. The last phase is for requirements management, in which changes to the requirements and reusable requirements are managed. These phases are carried out continuously until the end of the project.

The spiral line in Figure 3 represents an RE process loop. RE activities are cyclically executed throughout the project. In the educational workshop, the students elicit requirements, specify them, and evaluate them by conducting reviews with their clients. The educational aim is for the participants to not only to gain experience, but also to better understand the risks involved with starting the development base of their SRS. For example, if the SRS is developed based on only their first definition without sufficient knowledge of the domain and of their clients' business situation, they risk project failure. In particular, a cyclic evaluation is indispensable when the engineers do not have sufficient knowledge of their clients' domain or business.

As mentioned, our course is limited to two-and-a-half days. Since we do not have enough time to execute the whole cycle, we decided to set assumptions for the course. The course skips the third and last phases of the RE process based on the assumptions that professional engineers have prototyping experience and that there are tools for managing requirements.

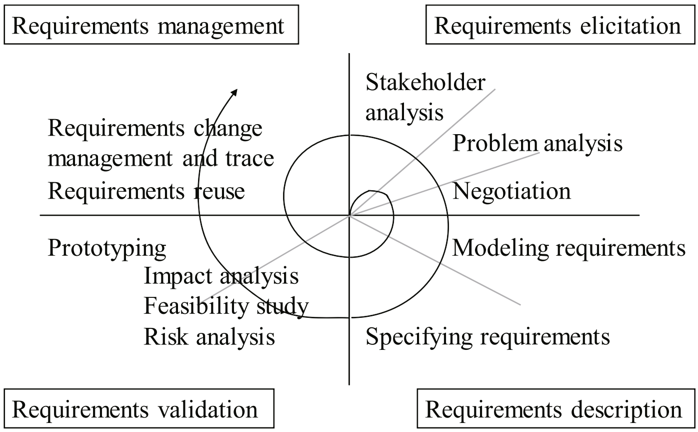


Fig. 3. RE process

One weak point often seen in engineers is that they tend to focus on solutions rather than on problems. This tendency leads them to take the client's demands at face value. As a result, when client requirements change, it might be too late for their project to succeed. Therefore, our target process has two steps: first, requirements elicitation and second, requirements modeling and SRS development. Throughout the course, the students learn how to gather information to write an SRS and experience how to develop a high-quality SRS.

3.2 Quality of SRS

When we started to design this course we had to determine what the goal of the course would be. The final product of the RE process is an SRS. As mentioned in Siemens' training course, engineers and their bosses expect our course to include high-quality SRS writing. Our course has to include a section on the IEEE standard 830-1998 (IEEE830 for short) in order to introduce the students to the goal of RE activities.

IEEE830 defines the quality characteristics of an SRS. An SRS should possess the following eight qualities (extracted from IEEE830).

- Correct:
  An SRS is correct if, and only if, every requirement stated therein is one that the software meets.
- Unambiguous:
  An SRS is unambiguous if, and only if, every requirement stated therein has only one interpretation.
- Complete:
  An SRS is complete if, and only if, it includes the following elements.
  - All the significant requirements.
  - A definition of the responses of the software to all the realizable classes of input data in all the realizable classes of given situations.
  - Full labels and references for all figures, tables, and diagrams in the SRS and a definition of all the terms and units of measure.

– Consistent:
  Consistency here refers to internal consistency. If an SRS does not agree with a higher-level document, it is not correct. An SRS is internally consistent if, and only if, no subset of individual requirements described in it conflicts. The three types of conflicts in an SRS are as follows:
  • The specified characteristics of real-world objects may conflict.
  • There may be logical or temporal conflicts between two specified actions.
  • Two or more requirements may describe the same real-world object but use different terms for that object.
– Ranked for importance and/or stability:
  An SRS is ranked for importance and/or stability if each requirement in it has an identifier to indicate either the importance or stability of that particular requirement.
– Verifiable:
  An SRS is verifiable if, and only if, every requirement stated therein is verifiable.
– Modifiable:
  An SRS is modifiable if, and only if, its structure and style are such that any changes to the requirements can be made easily, completely, and consistently while retaining the current structure and style.
– Traceable:
  An SRS is traceable if the origin of each of its requirements is clear and if it facilitates the referencing of each requirement in any future development or enhancement documentation.

Requirements engineers should understand these SRS quality characteristics and should be able to write high-quality SRSs. Professional engineers are motivated to produce good SRSs. Therefore, teachers of our course have to be able to explain the relationship between the course and the quality characteristics of SRSs.

3.3 Actual Project Situation for Requirements Changes

This section presents material for teaching engineers to visualize the types of requirements changes they face on the job. Requirements changes that cause a project to fail are a major problem for engineers. If the educational course teaches how to avoid unnecessary requirements changes, it may be possible to more effectively motivate the students.

There are development processes recommended for dealing with requirements changes. Many projects now follow incremental development processes so that new requirements and requirements changes can be incorporated as quickly as possible. The Agile Process [22] and the Unified Process [23] are two of the more popular examples of incremental development processes. They are also examples of an integrated RE process that integrates the RE process with other development processes [24]. Before learning these processes, professional engineers need to know the real characteristics of requirements changes.

Nakatani et al. [25] investigated the requirements processes of a project from its beginning to end. The engineers of the target project managed issue records, question

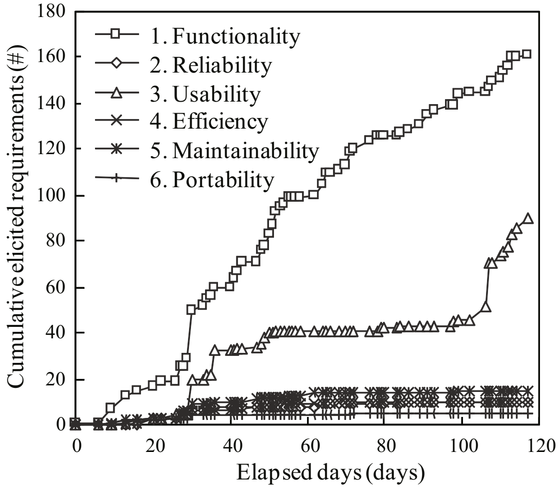


Fig. 4. Requirements changing process viewed from qualities viewpoint

and answer records, and decision records in each development phase. From these documents, they obtained not only the content of the requirements, but also each requirement's priority, the date of contact/agreement/completion, the name of the client's representative or manager, and the reasons for the change/addition/deletion. They analyzed the documents and counted the number of additional requirements, changed/adjusted requirements, and deleted requirements. Thus, if a new requirements was added, changed, and then deleted, it was counted as three requirements changes.

The requirements changing process viewed from a qualitative viewpoint is shown in Figure 4 in terms of the cumulative number of elicited requirements versus project days elapsed. The counted requirements changes were categorized into quality characteristics based on the ISO/IEC 9126 [26]. The functionality and usability requirements were elicited continuously throughout the project, while other requirements were frozen midway in the project schedule.

Figure 5 shows the requirements changing process from a component viewpoint. The counted requirements changes were categorized according to the components designed in the system. There were more than eight major components, some of which became relatively stable after the earliest stages of the project. On the other hand, some components were unstable. There are several factors that determine the degree of stability. If we can understand the situation surrounding the project, we can cope appropriately with the requirements changes [27]. Understanding the factors of requirements stability can help us prepare for requirements changes. Here are some examples of important project background information observed in the project and implications for future projects.

– Sufficient domain knowledge
  The continuous elicitation process of the functional requirements in Figure 4 represents the learning curve of the engineers and their clients. If the engineers already have a sufficient level of domain knowledge, they do not need to spend their time acquiring the knowledge. However, if they do not have enough domain knowledge,

they have to schedule time into their requirements elicitation process to learn the domain, otherwise they will face unexpected requirements changes caused by a lack of domain knowledge.

- Usability sensitivity
  Usability requirements were elicited throughout the project, but the project had no usability experts. Therefore, instability in the usability requirements was a continuous trial and error process. This shows that if the developing system is sensitive to its usability and the project does not have any usability experts, a lot of requirements changes to the user interface are expected. Thus, the project may adopt a trial and error process to achieve a better level of usability by accepting requirements changes.
- Market competitiveness
  Figure 5 shows two components that received a lot of requirements changes. These components were related to the market competitiveness of the product. Therefore, we can assume that if the developing system is placed in a competitive market, the requirements may be changed to follow market trends and/or keep market share. In such a situation, the susceptibility of each component to market trends should be ascertained. Then, the developers will be able to anticipate and cope with requirements changes due to market changes after the early stages of development.
- Non-functional requirements sensitivity
  The project was not affected much by non-functional requirements changes. However, engineers must be aware of non-functional requirements, since non-functional requirements affect the software architecture. Non-functional requirements can be derived from soft goals [28]. Therefore, the engineers have to master a method to deal with soft goals in order to maintain the software architecture.

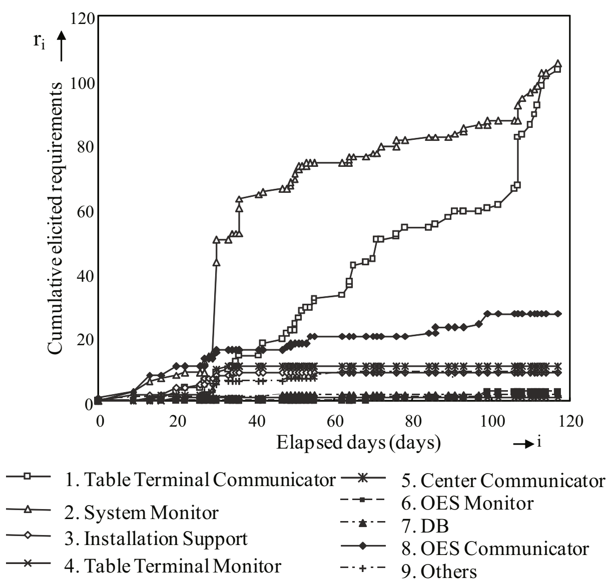


Fig. 5. Requirements changing process from component viewpoint

This case did not include any changes of mind, or "inconstancy," which sometimes causes serious problems. Fortunately, the project's clients were very cooperative. Engineers cannot usually expect such an ideal situation. If there are other effective and/or efficient solutions to the defined requirements, the engineers should carefully consider the alternatives to the requirements. Thus, the requirements engineers are responsible for defining any alternative requirements, as well as the selected ones. If they can explain the pros and cons of these requirements, their project can avoid groundless and/or unnecessary requirements changes.

The success of a project is influenced by the requirements engineer's ability to grasp the causes of the requirements changes. Goal-oriented methods can help an engineer visualize any alternative solutions and/or conflicts among the stakeholders and help them determine any possibilities for requirements changes. If the engineers can identify the stability of each requirement, then the designers who receive the requirements specifications can select an appropriate software architecture and apply the design patterns [29] in order to easily cope with future requirements changes. We can conclude from the research that learning goal-oriented analysis methods is essential for professional engineers.

3.4 Means of Achieving RE Goals

Then, we asked ourselves what kind of techniques should be taught to achieve this course's goal of teaching participants to write high-quality SRSs.

According to a report [30], less-experienced analysts often select a technique based on their experience. For example, the selected technique may be the only one they know, or they may think that a technique that worked well in the past is also appropriate this time. However, experienced engineers must consider the situational characteristics when they choose a technique. Furthermore, they need to know a set of additional predicates, which if true, cause experts to alter their primary choice. However, it is impossible to master the RE techniques at once.

There are a lot of RE techniques for eliciting, specifying, and managing requirements. Engineers have to apply them appropriately to solve problems. Tsumaki & Tamai [31] focused on two dimensions in order to characterize the variety of techniques: one concerning the elicitation of operational types and the other concerning the target object types. They divided the operational-type dimensions of the RE map into two categories, i.e., static and dynamic, and also divided the object-type dimension into closed and open types.

- The first dimension, the operational type, captures how the requirements elicitation and acquisition processes are conducted.
  1. Static
     Requirements are collected and sorted in a structured and systematic way. Generally, techniques of this type are suitable for dealing with a domain with a static structure. There are rules or guidance in decomposing/composing or classifying/clustering requirements related to objects, including domains/subdomains, goals, and the requirements themselves.

2. Dynamic
   Requirements are elicited in an imaginative and unmethodical way. Techniques of this type are suitable for dealing with a domain within its dynamic context. For many stakeholders, it is often easier to think in terms of procedural behaviors or situational changes over time. Requirements can be enumerated or generated through the imaginations of stakeholders or engineers. However, the use of such techniques is not limited to dynamic domains. For example, brainstorming is a typical technique of this type, but it can also be conveniently used for eliciting static or dynamic requirements.

– The second dimension, the object type, concerns the properties of the target object space to be analyzed, either relatively closed or open.
  1. Closed
     The object space is relatively stable, known, and closed. It can be understood by focusing mainly on the forms or syntax, because the space is basically structured and bounded by various kinds of syntactic forms.
  2. Open
     The object space is relatively unstable, unknown, changing, and open. The meanings have to be thoroughly considered in order to explore the space.

Various RE techniques can be mapped on this plane. Tsumaki & Tamai call this plane the RE technology map or RE map for short. The RE map allocated using typical techniques is outlined in Figure 6.

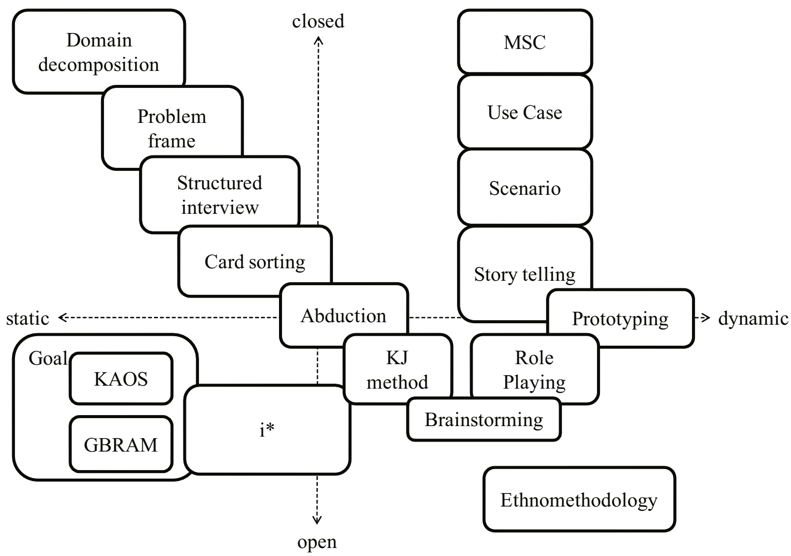


Fig. 6. RE technology map [31]

The following techniques are examples placed on the RE Map.

1. A sequence of scenario-based techniques can be observed in the right half. From top to bottom, they are message sequence chart (MSC), use case [32], scenario, story telling, and role playing. They share eliciting requirements properties based on experience and the user's or specialist's intuition in the progress of dynamic events, but differ in formality. The techniques at the top are relatively suitable for investigating closed spaces in a form- or syntax-conscious manner, while those at the bottom are suitable for investigating open spaces, thus tending to be more human-oriented. Also related to these techniques is the prototyping approach. Its major objective is to simulate dynamic interactions.
2. A sequence of knowledge-acquisition and classification-type approaches can be observed along the diagonal from the upper left to the lower right: domain decomposition, structured interview, card sorting, abduction, the KJ method [33], brainstorming, and ethnomethodology. The diagonal direction can be interpreted as the convergent/divergent (upper-left to lower-right) axis.
3. Of the goal-oriented approaches, KAOS [34] is formal and thus positioned higher, while the more informal GBRAM [35] is positioned lower. The i* framework [36] overlaps the goal-oriented approach, but extends to the right. There have been many proposals that have united the goal-oriented and scenario-based approaches [37]. They bridge the lower-left quadrant to the upper-right quadrant, and are not shown in the RE map.
4. As previously mentioned, both the top-down and bottom-up approaches of decomposing/composing entities have been included in the upper-left quadrant. The KJ method involves a process of clustering cards produced through brainstorming and is thus located closer to the upper-left hand side. The problem frame approach [38] is regarded as another way of decomposing the problem domain, although the decomposition is not structured with a top-down hierarchy, but is in parallel juxtaposition.
5. Various idea generation techniques can be included in the lower-right quadrant. For example, abduction-based approaches such as that by [39] can be allocated here.

Some methods combine techniques classified in different quadrants. For example, multiple viewpoint approaches can be positioned in the static-closed quadrant because they elicit requirements based on the classification of various viewpoints but they are usually combined with scenarios, goals, or other techniques to handle requirements from numerous viewpoints [40].

The course targets the requirements elicitation step: understanding the domain, identifying stakeholders, understanding stakeholders' problems and goals, deriving adequate solutions from the problems in order to achieve the goals, and helping negotiation. We can search for methods in the RE map for these five activities according to their object and operation types.

1. Understanding the domain: the operation type for understanding the domain is dynamic and its object type is open.
2. Identifying stakeholders: the operation type for identifying stakeholders is dynamic and its object type is open.
3. Understanding stakeholders' problems and goals: the operation type for understanding stakeholders' problems is relatively static and its object type is still open.

4. Deriving adequate solutions: the operation type for deriving adequate solutions is dynamic and its object type is closed.
5. Helping negotiation: the operation type for helping negotiation is dynamic and its object type is closed.

Engineers progress their process from the lower-right quadrant to the upper-right quadrant via the lower-left quadrant. In such a process, the upper-left quadrant is not focused upon. We can skip this quadrant, since the students learn the decomposition/composition way of thinking in the UML training course that precedes our course.

4 Instructional Design of Course

The purpose of the course is to teach participants the importance of the following.

- Understanding that doing RE activities is the only way to understand the background, purpose, and rationale of the requirements.
- Grasping the situation of stakeholders in order to extract the real world problems.
- Exploring alternative solutions and contribution relationships among the solutions and goals for starting negotiations.
- Identifying the variability and commonality of the requirements, as well as the likelihood of requirements changes for properly designing software architecture.
- Defining adequate requirements that can achieve goals without unexpected side effects in the real world.

Every RE technique addresses some part of the above list. In this section, we present the instructional design of our course, including the products, methods, educational style, and process.

4.1 Course Products

The product of RE is an SRS. RODAN [41,42], an integrated RE method [43], has the metamodel shown in Figure 7. The metamodel represents the structure of a requirement in the context of the real world. It has been developed from the product perspective of the RE activities. It provides the following information that an engineer must grasp in order to understand the requirements.

1. Each stakeholder's role and responsibilities in the organization.
2. Each stakeholder's interests, opinions, and world-views.
3. Each stakeholder's problems viewed from the perspective of his or her interests and/or responsibilities.
4. Achievable goals that could solve the problems.
5. Problematic as-is situations and expected to-be situations.
6. Solutions with alternatives to achieve each goal.
7. Using constraints to verify the selected solutions.

A business domain can have many roles. We can interpret the metamodel as starting from the business domain. In the following interpretation of the metamodel, we use italics for designating the "className" and *aClassName* or *className*s represent an instance or instances of the *ClassName*.

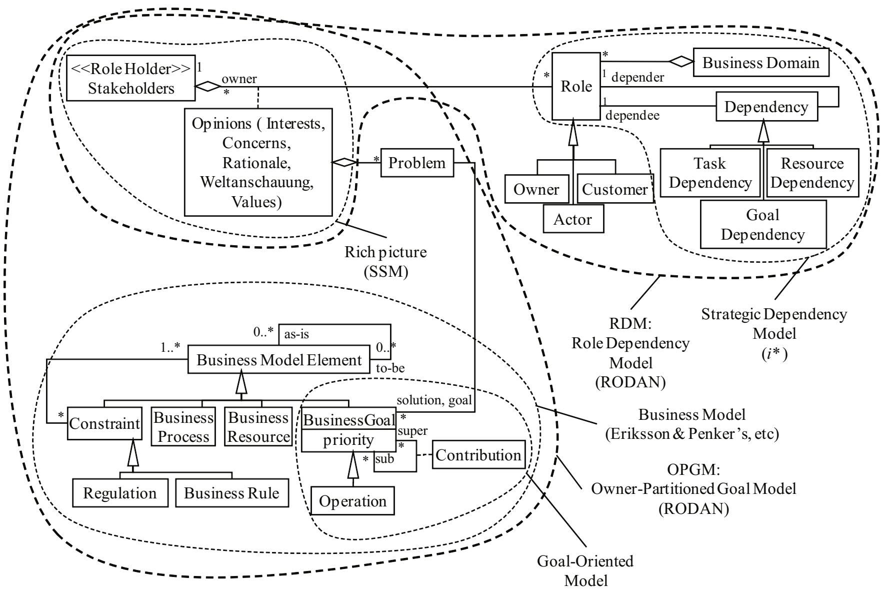


Fig. 7. Metamodel in RODAN from product perspective [42] with UML

1. Each stakeholder's role and responsibilities in the business domain.
   - A business domain (top right corner of Figure 7) aggregates various roles that are possessed by the stakeholders.
   - A role has dependencies with other roles in the business domain.
   - There are three kinds of roles in the business domain: owner, actor, and client [44].
   - Dependencies between roles are categorized into three dependencies: goal, task, and resource dependencies [36,45].

   Thus, *aStakeholder* owns multiple *role*s and those *role*s are aggregated by *aBusinessDomain*. This type of stakeholder is called a role holder in the task ontology [46]. *GoalDependency*, *TaskDependency*, and *ResourceDependency* are defined as the subclasses of class *Dependency*. *Owner*, *Actor*, and *Customer* are subclasses of class *Role*. For example, we can interpret that the goal of an owner role depends on the goal of the customer role, and the goal of an actor role depends on the goal of the owner role. We can give other examples of task dependency and resource dependency. An owner of a task directs the task to an actor in order to satisfy a customer, or an actor provides a resource owned by an owner to a customer.
2. Each stakeholder's interests, opinions, and world-views.
   - A stakeholder formulates an opinion according to his or her role.
   - The opinion is formed based upon interests, concerns, rationale, and world-views.

     A stakeholder's opinion represents what the stakeholder focuses on. It also represents his or her world-views based on his or her own roles.

   Thus, *Opinion* can be defined as an association class between *Stakeholder* and *Role*. The structure of *Opinion* implies that if the stakeholder's role changes, his or her opinion and world-view might also change.

3. Each stakeholder's problems viewed from his or her interests and/or responsibilities.
   - A stakeholder focuses on his or her problems through his or her opinions, interests, and world-views.

     Stakeholders recognize problems according to their responsibilities. If the view is transformed into another view, the problem changes into another problem. For example, if a person is responsible for delivering items, he or she may focus on the productivity of the delivery. However, if his or her interests become focused on managing the quality of items, his or her problems would be redefined.

   Thus, *anOpinion* has links with *problems*.
4. Achievable goals for solving problems.
   - A problem arises in the stakeholder's mind when he or she has difficulty achieving goals.

     Goals provide solutions or simple goals for the related problem. Operationalizing a goal derives the necessary operation. An operation for achieving a goal is a solution to the problem.
   - There are contribution relationships among the goals.

     Some goals and operations may contribute to achieving other goals, and others may not. There are contribution relationships among the goals. Those goals have to be prioritized by negotiation with the stakeholders, and then the goals and/or requirements for the developing system are selected.

   Thus, *Operation* is a subclass of *BusinessGoal* and there is a relationship between *BusinessGoal* and *Problem*. *BusinessGoal* has an attribute named "priority."
5. A problematic as-is situation and an expected to-be situation.

   An as-is situation and a to-be situation can be modeled using the business model elements of constraint, business process, business resource, and business goal [47]. *Constraint*, *BusinessProcess*, *BusinessResource*, and *BusinessGoal* are subclasses of *BusinessModelElement*.
6. Using constraints to verify solutions.

   A constraint is a business model element. Constraints for business processes and/or business resources are visualized in the as-is and to-be business models. If the selected requirements do not hurt the real world constraints, they can be accepted as correct requirements.

The metamodel shows that simply understanding a stakeholder's problem is not enough. The context of a problem, i.e., the opinion and/or world-view according to the stakeholder's role, must also be understood. Requirements are explained in the expected to-be situation, and operations can be concrete requirements.

The metamodel also shows the relation between the cause and effect of requirements changes. Requirements engineers have to distinguish stable roles from unstable roles in order to cope with requirements changes. For example, we can see that a role can cause requirements changes. A stakeholder recognizes a problem according to his or her opinion based on his or her role. A goal is defined to solve the problem, and requirements are derived to achieve the goal. The goal can change, if the role of the stakeholder changes.

According to the metamodel, the software requirements can trace their rationale backward via solutions that are included in the to-be business model. Furthermore, the solutions can be traced to their rationales backward via the goals, problems, stakeholder's value systems, and role dependencies. This is the backward traceability idea in RODAN.

Thus, the forward tracing path from the stakeholders and their roles to the solutions via their problems, opinions, and goals is regarded as the rationale of the requirements. The path shows how the software requirements affect the stakeholder's business.

Unselected solutions represent the alternatives to the candidate solutions. If the solution has a lot of alternatives, the requirements derived from the solution are regarded as a *hot spot* [48] in the requirements. This is important for designers because they are responsible for designing a robust software architecture that is not broken by requirements changes.

Again, the purpose of our course is to help engineers understand the importance of RE activities and motivate them to learn the RE techniques for grasping the stakeholders' situations. Students are expected to produce a model to visualize the requirements and their rationale. If they can understand the causes and effects of the requirements changes from the model, they will have motivation to learn RE and its technologies in order to prevent requirements changes to their projects. The model is also effective for cooperative works. If problems of a stakeholder are visualized in a model, the stakeholder can share them. Understanding and sharing others' problematic situations is the negotiation start point. Those requirements elements can be partially analyzed and modeled by existing methods. In the next section, we discuss the how we selected the methods for the course.

4.2 Method Selection

The course focuses on the importance of understanding real-world problems. To understand the real world, it is essential to integrate a business analysis and RE [49].

RODAN uses an integrated method to provide complete elements. Each dotted line in Figure 7 represents the structure of a model produced by a commonly used RE method. In RODAN, several methods are arranged to fit the metamodel: a role dependency analysis method that arranges i* with three different roles, i.e. actor, owner, and customer, and an owner-partitioned goal (OPG) analysis method arranged by KAOS [50] from the goal owner's viewpoint. These products have to be defined by a requirements engineer and they visualize what the engineer understands. Therefore, the course should cover the metamodel.

When we selected the methods for our course, we placed the products of each method in the metamodel. The metamodel is one of our guides for selecting methods to teach in the education course.

Furthermore, if we have the following evidential information, the SRS is more reliable.

1. Who needs the solutions and why? (Personal perspective)
2. Why should the solutions be highly prioritized? (Organizational perspective)

In the first analysis step, students think about the situation or business background. They have to focus on the open-type objects and use dynamic methods. The applicable methods are in the lower-right quadrant of the RE map.

In the second step, the students have to think about solutions that could solve the problematic situation or achieve the business goals. Here, students can use the static methods, e.g., the goal-oriented methods, in the lower-left quadrant. The target object type is still open since the activities require imagination. For example, when a student uses KAOS, he or she develops goal model by imaging various situations in order to achieve the goals and also derives alternative subgoals. These goals and subgoals including operationalized solutions are represented in the goal model. Finally, he or she has a full understanding of the selectable requirements or solutions in the goal model. Then, the target object type becomes closed.

The next consideration is an evaluation of the adequacy of the requirements. Software is in constant danger of attacks by malicious people. Furthermore, software systems affect the real world and sometimes behave differently than users expect. Students of the course should analyze the possible side effects and/or malicious behaviors that might harm users in the real world. In these analyses, the target object type is closed and the applied methods require the imagination of analysts [51]. These methods are positioned in the upper-right quadrant of the RE map. Scenario analysis, claim analysis [52], and misuse case analysis [53,54] are the applicable methods.

After eliciting the requirements, students must consider the appropriateness of the software. These are static activities, and they treat closed-type objects. For example, the model-checking approach is an applicable method for one of these activities.

The real requirements elicitation process starts from various situations. In some cases, analysts have to start by understanding a vague, real-world problem. In such a situation, the analyst proceeds with analysis from the lower-right quadrant to the lower-left quadrant, and then goes to the upper-left quadrant via the upper-right quadrant. In other situations, the analyst can start by verifying the correctness of the provided requirements, starting from the upper-left quadrant.

Our educational course is designed on the assumption that the target domain is vague. We have to keep in mind that the purpose of this course is to teach the importance of RE so that engineers become aware of the stability of the requirements. We expect the students to be able to adjust the requirements process to the project situation and apply their own knowledge after the course.

We selected the methods used in the course based on the RE map and the RODAN metamodel. To fit the existing methods into the metamodel, we extend these methods.

- Rich picture:
  This is a tool introduced in the Soft Systems Methodology (SSM) by Checkland [44]. In a rich picture, a balloon represents stakeholders opinion of the problematic situation. The place of an icon of each stakeholder can represent environment with other stakeholders. From the expression of the icon, we can read the stakeholder's subjective emotion. Thus, rich pictures are useful for depicting a stakeholder's situation or environment. They treat open-type objects dynamically. Therefore, they can be positioned to the upper left position of the ethnomethodology [55] and real-world interview and observation in the lower-right quadrant of the RE map.

- CATWOE analysis:
  This is another tool introduced in SSM. CATWOE[1] is interpreted as, "the *actor* is ordered by the *owner* to *transform* the present problematic situation into an accomplishable future situation for his or her *customers* in a given *environment*," and represents an owner's intention for the future. It uses two kinds of analyses. The "CAT" and "OE" are dynamically derived from open type objects. Therefore, they can be positioned to the upper left of rich picture analysis in the lower-right quadrant of the RE map. On the other hand, the "W" that represents the *owner*'s world-view and his or her perspectives are analyzed by abductive reasoning. It can be positioned in the same spot as abduction.
- Role dependency modeling:
  This modeling technique extends the strategy dependency graph in i* [36] with the dependency roles. Therefore, it can be positioned in the same spot as i*.
- Owner-partitioned goal analysis:
  This technique extends the goal model of KAOS [56] partitioned by the goal owners. Thus, it can be positioned in the same spot as KAOS. This technique is used for negotiation and is placed in the lower-left quadrant.
- Claim analysis:
  This technique is introduced in the scenario analysis [52] in the upper-right quadrant.
- Misuse case analysis:
  It is positioned as part of the use case analysis in the upper-right quadrant.
- Object-oriented methods:
  There are various modeling techniques with the Unified Modeling Language (UML) [57]. It is essential to have a UML technical course as a preceding course, since the RE course does not have enough time to teach these techniques. The object-oriented methods are useful for analyzing business world, e.g., interaction across organizations, the structures of organizations, and business resources [47]. They can be positioned in the upper right quadrant.

Now, we are ready to define the RE educational map used in our course (Figure 8). The gray rectangles represent methods that are mentioned in the lecture.

4.3 Style Selection

As mentioned in the first section, our target students are engineers with over ten years of experience. Our course focuses on the process of eliciting requirements. Since most professional engineers are not aware of their clients' mindsets, the students have to recall them. This is the key when they interview their actual clients. Therefore, all students need to have a chance in the course to experience the client's role.

Beus-Dukic et al. mentioned that the teaching of modeling methods is not straight forward [54]. A workshop-style course can provide a chance for students to work through a lot of problems and exercise modeling methods through trial and error.

[1] CATWOE is an abbreviation of *customer*, *actor*, *transformation process*, *world view*, *owner*, and *environment*.

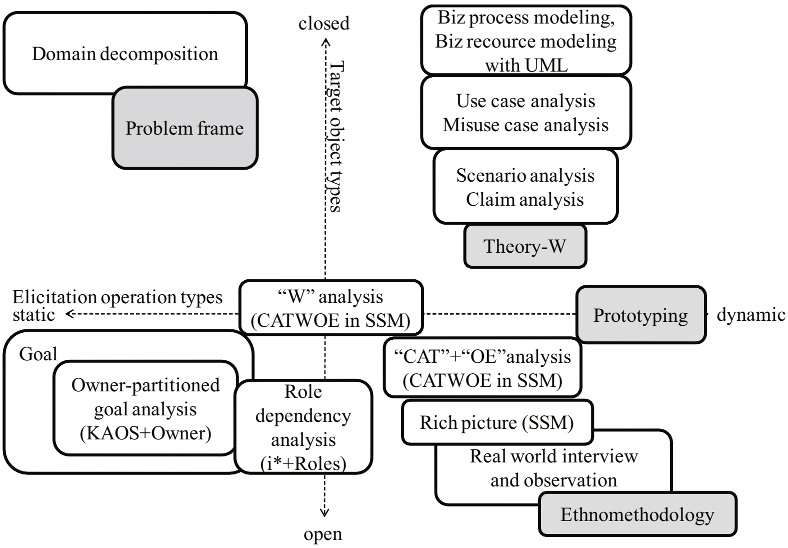


Fig. 8. RE technology map for education

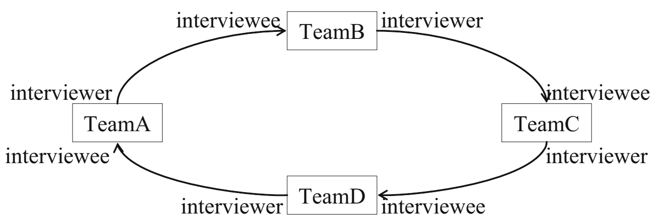


Fig. 9. Example of course organization

We chose a role-playing workshop as our course style. Figure 9 shows an example of the course organization. Students from TeamA are the developers for TeamB, and they play the client role for TeamD. Once they interview for ten to fifteen minutes with TeamB, in the next session they switch to the client role to be interviewed by TeamD. The students are required to develop an SRS for their clients and then evaluate the SRS with their clients.

Each team consists of four to five students, which is a good size for effective discussions and cooperation. It is better if the class consists of more than two teams to prevent the students from interviewing with the interviewed team. Time management is more efficient with an even number of teams.

4.4 Schedule Planning

The course is designed for teaching the early RE process, i.e. stakeholder analysis, problem analysis, and negotiation. The course is limited to two and a half days.

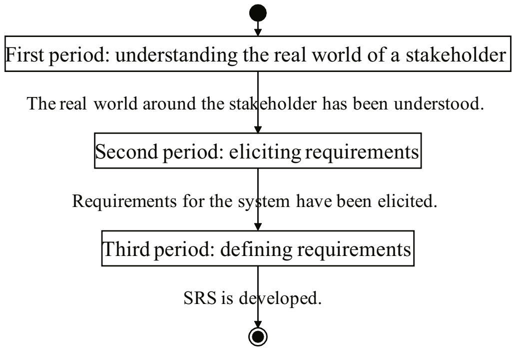


Fig. 10. Course activities and their exit criteria

In the course, students start by focusing on clients' personal viewpoints by drawing rich pictures. The activities of each educational step along with their exit criteria are shown in Figure 10. The course consists of three periods.

1. The first period (eight hours):
   After the first period, the students are expected to consider real-world problems in their models and use their models as evidence to answer the first question, "Who needs the solutions and why?" The details of the period are shown in Figure 11.
   The first aim of this period is to use a rich picture to learn the relationships between the stakeholders' viewpoints and perspectives.
   The second aim is to extract the organizational goals and requirements using CATWOE and role dependency (RD) analyses. The course then proceeds with a lecture, exercise, and evaluation cycle.

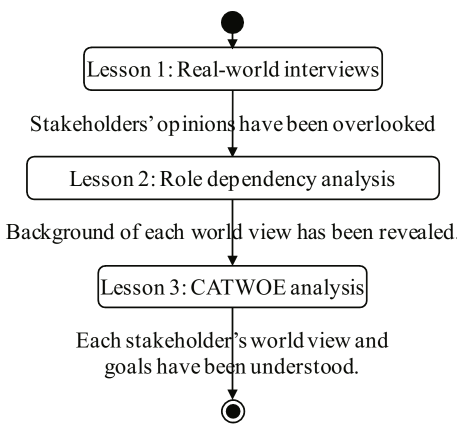


Fig. 11. Activities in first period

- Lesson 1 (four hours):
  - Lecture: Introduction to RE, IEEE830, and rich pictures.
  - Exercise: Draw rich pictures to grasp the real world situation from the viewpoint of the stakeholders' by interviewing them.
  - Evaluation: Review the rich pictures with stakeholders.
- Lesson 2 (two hours):
  - Lecture: Introduction to RD analysis. The RD model is presented as a model of "the owner ordering the actor to do something for the customers."
  - Exercise: Perform an RD analysis and extract stable organizational goals and their owners.
  - Evaluation: Review the RD models with the stakeholders.
- Lesson 3 (two hours):
  - Lecture: Introduction to CATWOE analysis.
  - Exercise: Define CATWOE.
  - Evaluation: Review and validate the results of the analysis with the stakeholders.

2. The second period (six hours):
   At the end of the second period, students should be able to answer the questions "Why should the solutions be highly prioritized?" and "Why should we directly interview multiple stakeholders?" The details for this period are shown in Figure 12.

   The first aim of this period is for the students to solve organizational problems using owner-positioned goal (OPG) analysis to extract the requirements and their alternatives.

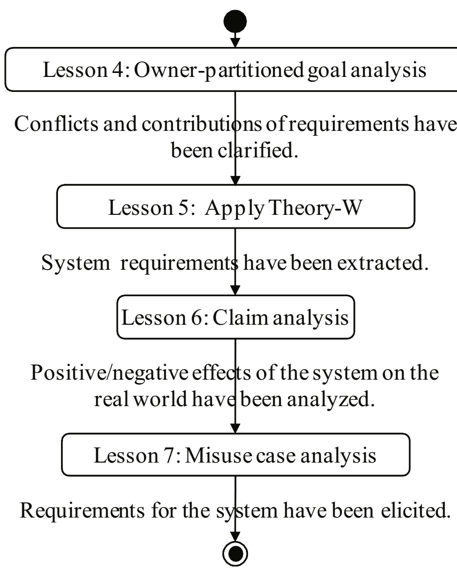


Fig. 12. Activities in second period

The second aim is for the students to understand the rationale of the requirements from the CATWOE definitions and the owners' goals. The students are also required to represent the as-is and to-be situations from structural, dynamic, and collaborative perspectives with UML. The as-is models represent the real problematic situations and the to-be models represent the accomplishable future situations.

The third aim is to improve the students' ability to evaluate quality of the requirements by considering misuse cases and a scenario analysis.

   - Lesson 4 (one hour):
     - Lecture: Introduction to OPG analysis.
     - Exercise: Develop an OPG model.
     - Evaluation: Review OPG models with the stakeholders.
   - Lesson 5 (one hour):
     - Lecture: Introduction to Theory-W as an example of a negotiation technique.
     - Exercise: Elicit the system requirements using OPG models to develop win-win situations with the stakeholders.
     - Evaluation: Review the system requirements with the stakeholders.
   - Lesson 6 (two hours):
     - Lecture: UML and scenario analysis.
     - Exercise: Analyze the as-is and to-be situations with UML and apply the scenario analysis to validate the requirements for supporting the proper stakeholders.
     - Evaluation: Review the adequacy of the requirements and whether they affect the stakeholders positively or negatively.
   - Lesson 7 (two hours):
     - Lecture: Introduction to misuse case analysis to protect the system from malicious actors.
     - Exercise: Develop use case diagrams for misuse cases.
     - Evaluation: Review the acceptability of the requirements and whether they can prevent the system from the intentions of negative stakeholders.
3. The third period (four hours):
   Each team develops an SRS. The details of the period are shown in Figure 13.
   - Lecture: Review of IEEE830. Review a sample SRS, rather than teaching how to write SRSs.
   - Exercise: Develop an SRS from the models developed in the workshop.
   - Evaluation: Review SRS with the stakeholders.

5 Course Experience

5.1 Workshop Assignment

We have conducted the course 16 times in the past four years. It was usually run with three to six teams. Each team was composed of four to five engineers with more than ten years of software development experience. The following assignment was used in the course:

> *Develop a software requirements specification for an effective meeting support system.*

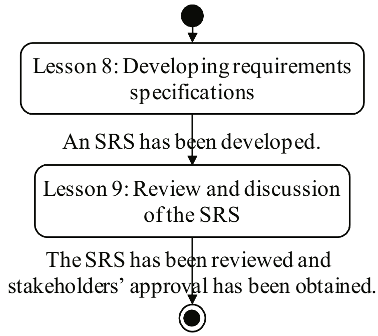


Fig. 13. Activities in third period

Vagueness is important for achieving the goals of the course. The assignment description is sufficiently vague. Students have to consider what it means for meetings to be effective or ineffective, as well as what it means to support meetings. For the course evaluation, we used "learned memos (LM)", which are shared post-it notes that list what each student learned. The students made LMs after each review session and put them up on the wall. LMs tell us what the students learned and when they learned it.

Figure 14 shows a review session, the LMs, and a rich picture from a course. In a review, interview team members observe their clients and read their faces to monitor the process of the review. While one member explains a product, the other members stand beside him and observe their clients' faces. If the explanation is not understandable, they may maintain a straight face or have a hardened face. It is easy for professional

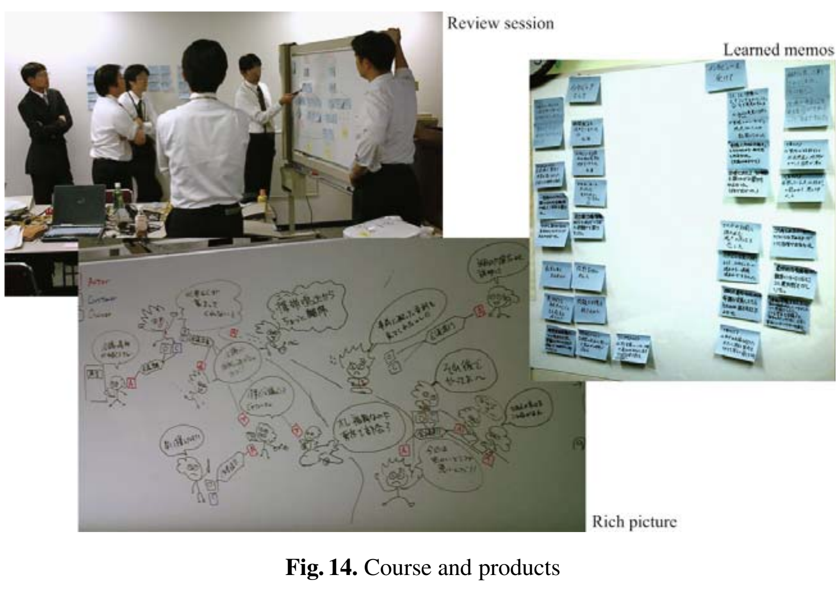


Fig. 14. Course and products

engineers to explain their products, however, it is hard to observe someone's thoughts. Expressions reveal unspoken information. Requirements analysts should never neglect this important resource.

5.2 Products

In a workshop, the requirements elicitation started from the interviews conducted through role-playing. Students playing the roles of engineers drew a representation of the stakeholders' opinions in a rich picture. Then, the team of analysts defined the role dependencies in the picture as *actor*, *customer*, and *owner*. Thus, the rich picture became the RD model.

An example of an RD model is depicted in Figure 15. There were various people in the meeting domain. The teachers guided the students to help them grasp the stakeholders' world-views.

The CATWOE lesson is effective for understanding the business domain. In the interview, some stakeholders may not clearly describe *W* and *O*, but do well with *C*, *A*, and *T*. To define the "who" and "why" aspects of the requirements, the students need to grasp the *O* and *W* for each stakeholder's opinion. The CATWOE lesson is an exercise in recognizing how many different interpretations there can be for the same "C," "A," and "T." The following CATWOE [44] examples are derived from Figure 15. They are defined for the same *C*, *A*, *T*, and *E*, but with different owners. The definitions represent each *W* as being different for each *O*.

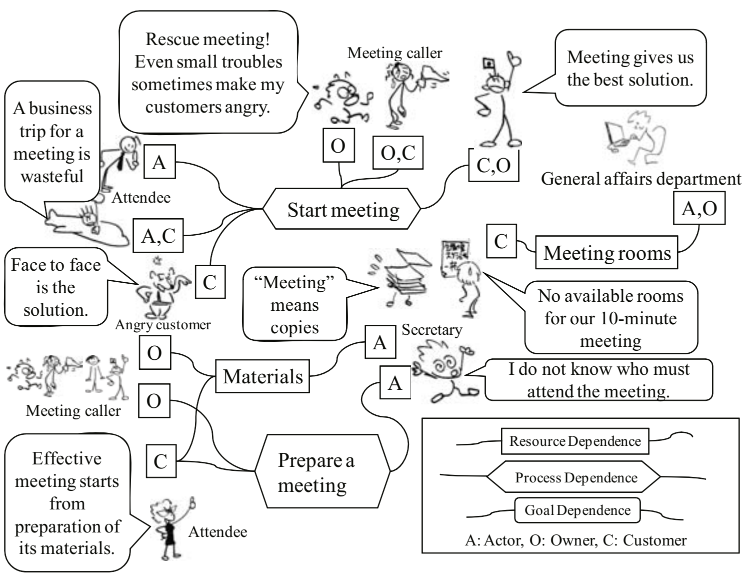


Fig. 15. Product example

- *Customer*: The person who wants to call a 10-minute meeting.
- *Actor*: The person who arranges the meeting.
- *Transformation process*:
  The situation "a short meeting cannot be held" is transformed into the future situation: "a short meeting can be held."
- *Environment*: Meeting rooms are available for short meetings.

- Case 1
  - *Owner*: The person who wants to call a 10-minute meeting.
  - *World view*: Short meetings help us understand each other more smoothly and help enable us to make quick decisions that improve business productivity.
- Case 2
  - *Owner*: The person who arranges the meeting.
  - *World view*: I don't expect my boss to assess me as incompetent because of a problem with the meeting room scheduling.

After the students defined these CATWOEs, their clients selected one of the CATWOEs and/or corrected the misunderstood part. The students can learn through these exercises that RE should be practiced cooperatively by analysts and stakeholders.

The defined CATWOE can be interpreted in a composed structure according to the why, what, and how information. For example, the *W* is "why" for the execution of *T*, and the execution of *T* is "what" for the *E*. Thus, by the end of the CATWOE exercise, the workshop was ready to begin a goal-oriented analysis. These three items were used as the nodes of a goal tree.

Goal-oriented modeling provides several leaves for the root goals. The students could recognize the openness of the solution domain because of the positions of the goal-oriented methods in the RE map. Requirements triage [58] and win-win negotiation [59] would be conducted in the workshop.

5.3 Evaluation

There are several evaluation methods that can be used for training. There are five levels of evaluation, including what Jack Phillips [60] defined as return on investment (ROI) methodology.

1. Reaction and/or satisfaction and planned action:
   Measures of participant satisfaction with the training program.
2. Learning:
   Measures of participants' changes in knowledge, skills, and attitudes during the training.
3. Job applications and/or implementation:
   Measures of participants' changes in the on-the-job behavior and progress using what they have learned from the training.
4. Business impact:
   Measures the actual business results achieved as a consequence of applying the knowledge and skills from the training.
5. Return on investment (ROI):
   Compares the program's monetary benefits to the actual cost of the training program.

The first and second level evaluations were conducted by a simple questionnaire and by asking for the students' candid comments in the second trial year, which was 2007. The questionnaire is composed of four questions scored on a scale of 1 to 5 for evaluating the effectiveness of the course for their job, its understandability, the teacher's ability, and the material quality. There were twelve students in three groups with average scores of 3.70, 3.00, 3.80, and 3.60 out of 5, respectively. Although the scores were not so high, according to the course sponsor, the scores depended on the student's business background. For example, there were some students who did not have any experience with requirements modeling. When students felt the modeling was difficult, it is more difficult for them to start learning RE.

The course was planned under the assumption that the students had already mastered UML. However, a person may master UML without knowing the modeling technique. To improve the effectiveness of our course, we rescheduled the RE related courses in 2008 to come after the a course on modeling with UML. In the modeling course, the students discussed how to create appropriate models and learned the key to successful modeling, namely, that there are no correct answers, but there are adequate ones. As a result, the average of the metrics became 4.20, 3.67, 4.07, and 3.87. Thus, we improved the results.

One of the reasons why the understandability scores were low was explained in the comments section of the questionnaire. Although the students seemed to understand the importance of requirements engineering, they felt they needed more experience to fully understand the methods. The purpose of the course is to motivate the students to start learning the RE methods. Therefore, this comment was expected from the students.

In one of the courses that had 15 students, we collected 67 and 70 LMs as comments about the stakeholder and analyst roles, respectively. Furthermore, there were 28 LMs as RODAN evaluations and 6 LMs for the SRS as a product of the exercise. Representative comments from the students are given below. The following evaluation results were collected from the students' learning memos.

1. What students learned.
   - There were unproductive questions:
     When a student with a client role was asked by another student with an engineer role, "This is our understanding of your intention. Is it ok?" he could not reply, "no," even though the model did not completely represent his situation. He felt that the question, "Does it correctly represent your intention?" would be a better question. He learned though his client role experience that the initial question might confuse a stakeholder.
   - Requirements are not for the software engineers, but for the clients:
     One particular student found that the requirements were not the product for an engineer, but for the stakeholders. She also mentioned that she had to offer solutions that the stakeholders could accept rather than ones that the she wanted to build. In another class, when the developer team presented their SRS to their client team, the client team asked "so, what does the system work for our problems?" Analyzing and providing solutions "for the clients" is one of the keys of RE.

- A solution is not a solution for everyone:
  One of the students realized that there were different problems and solutions for every stakeholder and mentioned that one solution could not solve everyone's problems. The course instruction is designed according to the RODAN metamodel. This model shows that every requirement depends on the stakeholder's role. We were able to guide students with the RE methods to see the world that the metamodel implies.
- Requirements need rationale:
  Several students in the client role mentioned that, "Some of the requirements in the SRS did not have rationales. When the requirements were decomposed into detailed functional requirements, their rationales became vague." Then, they had to re-explore the rationale of the requirements from their exercise products in the final exercise. Interestingly, when the analyst team presented their SRS in the final exercise, the stakeholder team presented the following question: "Then, what goals will be achieved by this system? How does this system change our world?" Although the analyst team members almost forgot the purpose of the course during their SRS development exercise, their client's team members did not forget it and pointed it out. We hope that such an experience will help students better master RE in their jobs.
- The rationale of requirements exists in the stakeholders' real-world:
  Another student commented that he understood the reason why stakeholders frequently change their requirements. Analysts should grasp the rationale of requirements from their clients' real world. The comment came up after the student played the client role. Role-playing style was effective for students to understand their clients mindsets.

Thus, the course succeeded in helping the students focus on the rationale behind the requirements. It also succeeded in teaching the importance of requirements engineering.

2. Strengths and weaknesses
The strength of the course is that it was developed for professional engineers. Therefore, we could assume that the students were well grounded in software engineering. We did not spend any time teaching basic software engineering techniques. However, we could not efficiently achieve our goals. The first two years told us that we assumed too much expertise for some students.

One novice engineer who was not familiar with static and dynamic modeling techniques needed an additional week of modeling training. UML modeling training was given before this course from the third year on. This improved the course results.

Although their technical experiences were different, the course assignments were familiar to them. The role-playing workshop was an effective exercise for getting them to join in and to directly experience being a stakeholder.

During the first period of the workshop, the students learn how many different viewpoints exist in a domain. Since the students were experienced engineers, we expected them to have faced requirements changes regularly. Therefore, it was not difficult to teach them the need to analyze the stakeholders' world-views in order to save their projects from unexpected requirements changes. However, the students in the course were selected and told to attend the course by their bosses. A problem arose from the variety of student backgrounds in the course. Some of them believed that they would have fewer opportunities to come in contact with their clients. These engineers usually start their jobs with an SRS provided by other engineers. It was hard to motivate them to start learning RE. The course may be effective only for professional engineers who directly communicate with their clients in their work.

6 Conclusions

There are various RE educational courses. Most of them are designed for undergraduate students. Our target students are engineers with over ten years of experience. They have programming, design, and communication skills, as well as team-work skills. The instruction is designed from the product and process perspectives, as well as the method application perspective. The course follows the RE process of requirements elicitation. The products of the course are designed using the RODAN metamodel. The metamodel represents the structure of a requirement with its rationale and background. The methods are selected to fit the metamodel and are positioned on the RE map from the lower-right quadrant to the upper-right quadrant via the lower-left quadrant. The course is laid out over two and a half days. We conclude from the learned memos that the course was well designed for enabling students to understand the importance of RE and for motivating them to start learning RE methods. After the course, the students understood the importance of understanding the background and/or rationale of the requirements for producing adequate SRSs.

The course introduces several methods, but these methods are not mandatory for all engineers. The metamodel of the products of the course can accept other methods if their products fit the metamodel and have a proper position on the RE map. Consequently, the students are responsible for selecting adequate methods and applying them as requirements elicitation methods. Such adaptability is important for professional engineers, because each software development company has its own standard development process and methods, including those for RE. Finding ways to integrate RE methods into future projects is the responsibility of the students themselves. They may use methods that they are familiar with instead of the SSM or OPG analysis and/or choose their methods in response to the actual situation.

Acknowledgements. We would like to thank all the students that took the course and also offer special thanks to the course staff for helping to improve our course.

Resources

A Key Books

- Checkland, P. and Scholes, J.: Soft Systems Methodology in Action. John Wiley & Sons (1991)
- Gause, D. C. and Weinberg, G. M.: Exploring Requirements: Quality Before Design. Dorset House (1989)
- Gause, D. C. and Weinberg, G. M.: Are Your Lights On?: How to Figure Out What the Problem Really Is. Dorset House (1990)
- Jackson, M.: Problem Frames. Addison-Wesley (2001)

- Jacobson, I., Christerson, M., Jonsson, P. and Overgaard, G.: Object-Oriented Software Engineering. Addison-Wesley (1992)
- Van Lamsweerde, A.: Requirements Engineering: From System Goals to UML Models to Software Specifications. Wiley (2009)
- Robertson, S. and Robertson, J.: Mastering the Requirements Process (2nd edition). Addison-Wesley (2006)
- Sommerville, I. and Sawyer, P.: Requirements Engineering: A Good Practice Guide. Wiley (1997)
- Kalle J. Lyytinen, K. J., Loucopoulos, P., Mylopoulos, J. and Robinson, W. (edit.): Design Requirements Engineering: A Ten-Year Perspective, (Lecture Notes in Business Information Processing) Springer (2009)

B Key Survey/Review Articles

- Berry, D.M. and Kamsties, E.: Chapter 2: Ambiguity in requirements specification. In: J.C.S.D.P. Leite, J. J. H. Doorn (eds.) Perspectives on Software Requirements, pp.7–44 (2004)
- Kruchten, P.: The 4+1 view model of architecture. IEEE Software 12(6), 42–50 (1995)
- Van Lamsweerde, A.: Goal-oriented requirements engineering: A roundtrip from research to practice. In: Proc. of the 12th International Requirements Engineering Conference (REf04), pp.4–7. IEEE (2004)
- Nuseibeh, B. and Easterbrook, S.: Requirements engineering: a roadmap. In: Proc. of the Conference on The Future of Software Engineering, pp.35–46. ACM (2000)
- Zave, P. and Jackson, M.: Four dark corners of requirements engineering. ACM Transactions on Software Engineering and Methodology 6(1), 1–30 (1997)

C Key Journals

- Requirements Engineering Journal, Springer London

D Key International Conferences/Workshops

- CAiSE–International Conference on Advanced Information Systems Engineering
- ICSE–International Conference on Software Engineering
- RE–IEEE International Requirements Engineering Conference
- REET–International Workshop on Requirements Engineering Education and Training

E Software

- Objectiver– Objectiver is a tool for KAOS and is downloadable from
  http://www.objectiver.com/
- *i** tools– OME is a tool for goal-oriented and/or agent-oriented modeling and analysis. It supports the *i**, NFR modeling. It is available at
  https://se.cs.toronto.edu/trac/ome
- UML tools– Many free tools for object-oriented analysis and modeling with UML are/will be available on the WWW.

References

1. Tsuji, H., Sakurai, A., Yoshida, K., Tiwana, A., Bush, A.: Questionnaire-based risk assessment scheme for Japanese offshore software outsourcing. In: Meyer, B., Joseph, M. (eds.) SEAFOOD 2007. LNCS, vol. 4716, pp. 114–127. Springer, Heidelberg (2007)
2. Tiwana, A., Bush, A.A., Tsuji, H., Yoshida, K., Sakurai, A.: Myths and paradoxes in Japanese IT offshoring. Commun. ACM 51(10), 141–145 (2008)
3. Sheng, Z., Tsuji, H., Sakurai, A., Yoshida, K., Nakatani, T.: Preliminary analysis for risk finding in offshore software outsourcing from vendor's viewpoint. In: Proc. of the Software Engineering Approaches for Offshore and Outsourced Development 2008, SEAFOOD 2008 (2008)
4. Boehm, B.W.: Software Engineering Economics. Prentice-Hall, Englewood Cliffs (1981)
5. Berenbach, B.: A hole in the curriculum. In: The 1st International Workshop on Requirements Engineering Education and Training, REET 2005 (2005), `http://re.cti.depaul.edu/REET05/P2_Berenbach.pdf`
6. Gause, D.: Making the case for the soft in software engineering or what does it really take to be prepared for the RE world ahead (2005), `http://lams.epfl.ch/course/esoa2008/old_esoa2007/business-it/esoa2007_week6b_gause.pdf`
7. Nikula, U.: Experiences of embedding training in a basic requirements engineering method. In: Proc. of the 17th Conference on Software Engineering Education & Training (CSEET 2004), pp. 104–109. IEEE, Los Alamitos (2004)
8. Lami, G.: Teaching requirements engineering in the small: an under-graduate course experience. In: The 1st International Workshop on Requirements Engineering Education and Training, REET 2005 (2005), `http://re.cti.depaul.edu/REET05/P4_Lami.pdf`
9. Berenbach, B., Wolf, T.: A unified requirements model; integrating features, use cases, requirements, requirements analysis and hazard analysis. In: Proc. of the International Conference on Global Software Engineering (ICGSE 2007), pp. 197–203. IEEE, Los Alamitos (2007)
10. Smith, R., Gotel, O.: Using a game to introduce lightweight requirements engineering. In: Proc. of the 15th International Requirements Engineering Conference (RE 2007), pp. 379–380. IEEE, Los Alamitos (2007)
11. Beatty, J., Alexander, M.: Games-based requirements engineering training: an initial experience report. In: Proc. of the 16th International Requirements Engineering Conference (RE 2008), pp. 211–216. IEEE, Los Alamitos (2008)
12. Zowghi, D., Paryani, S.: Teaching requirements engineering through role playing: Lessons learnt. In: Proc. of the 11th International Requirements Engineering Conference (RE 2003), pp. 233–241. IEEE, Los Alamitos (2003)
13. Regev, G., Gause, D.C., Wegmann, A.: Requirements engineering education in the 21st century, an experiential learning approach. In: Proc. of the 16th International Requirements Engineering Conference (RE 2008), pp. 85–94. IEEE, Los Alamitos (2008)
14. Madhavji, N.H., Miller, J.: Investigation-based requirements engineering education. In: The 1st International Workshop on Requirements Engineering Education and Training, REET 2005 (2005), `http://re.cti.depaul.edu/REET05/P3_Madjhavi.pdf`
15. Berenbach, B., Rayment, T.: The evaluation of a requirements engineering training program at Siemens. In: Proc. of the 16th International Requirements Engineering Conference (RE 2008), pp. 205–210 (2008)
16. Simmons, E.: Reflections on a successful corporate requirements engineering training curriculum. In: The 1st International Workshop on Requirements Engineering Education and Training, REET 2005 (2005), `http://re.cti.depaul.edu/REET05/F3_Simmons.pdf`
17. Wegmann, A.: Theory and practice behind the course designing enterprise-wide IT systems. IEEE Transactions on Education 47(4), 490–496 (2004)

18. Shaw, M.: Prospects for an engineering discipline of software. IEEE Software 07(6), 15–24 (1990)
19. Nuseibeh, B., Easterbrook, S.: Requirements engineering: a roadmap. In: Proc. of the Conference on The Future of Software Engineering, pp. 35–46. ACM, New York (2000)
20. IEEEstd.830-1998: Recommended Practice for Software Requirements Specifications. IEEE, Los Alamitos (1998)
21. Ohnishi, A.: Report of requirements engineering working group activities (in Japanese). IPSJ SIG Technical Reports, 2001-SE-130-18, 127–134. IPSJ (2001)
22. Beck, K., et al.: Manifesto for agile software development (2001), `http://agilemanifesto.org/`
23. Jacobson, I., Booch, G., Rumbaugh, J.: The Unified Software Development Process. Addison-Wesley, Reading (1999)
24. Sommerville, I.: Integrated requirements engineering: A tutorial. In: IEEE Software, pp. 16–23. IEEE, Los Alamitos (2005)
25. Nakatani, T., Hori, S., Ubayashi, N., Katamine, K., Hashimoto, M.: A case study: Requirements elicitation processes throughout a project. In: Proc. of the 16th International Requirements Engineering Conference (RE 2008), pp. 241–246. IEEE, Los Alamitos (2008)
26. ISO/IEC: 9126: Software engineering–product quality (2001)
27. Nakatani, T., Hori, S., Tsuda, M., Inoki, M., Katamine, K., Hashimoto, M.: Towards a strategic requirements elicitation: A proposal of the PRINCE model. In: Proc. of the 4th International Conference on Software and Data Technologies (ICSOFT 2009). INSTICC (2009)
28. Chung, L., Nixon, B.A., Yu, E., Mylopoulos, J.: Non-Functional Requirements in Software Engineering. Kluwer Academic Publishers, Dordrecht (1999)
29. Gamma, E., Helm, R., Johnson, R., Vlissides, J.: Design Patterns. Addison-Wesley, Reading (1995)
30. Hickey, A.M., Davis, A.M.: Elicitation technique selection: How do experts do it? In: Proc. of the 11th International Requirements Engineering Conference (RE 2003), pp. 169–178. IEEE, Los Alamitos (2003)
31. Tsumaki, T., Tamai, T.: A framework for matching requirements elicitation techniques to project characteristics. Software Process: Improvement and Practice 11(5), 505–519 (2006)
32. Jacobson, I., Christerson, M., Jonsson, P., Overgaard, G.: Object-Oriented Software Engineering. Addison-Wesley, Reading (1992)
33. Kawakita, J.: Hassouhou (in Japanese). Chuou Kouron Sha (1967)
34. van Lamsweerde, A.: Goal-oriented requirements engineering: A roundtrip from research to practice. In: Proc. of the 12th International Requirements Engineering Conference (RE 2004), pp. 4–7. IEEE, Los Alamitos (2004)
35. Anton, A.I.: Goal-based requirements analysis. In: Proc. of the Second International Conference on Requirements Engineering (ICRE 1996), pp. 136–144. IEEE, Los Alamitos (1996)
36. Yu, E.S.K.: Towards modelling and reasoning support for early-phase requirements engineering. In: Proc. of the 3rd International Symposium on Requirements Engineering (RE 1997), pp. 226–235. IEEE, Los Alamitos (1997)
37. Rolland, C., Souveyet, C., Achour, C.B.: Guiding goal modeling using scenarios. IEEE Transactions on Software Engineering 24(12), 1055–1071 (1998)
38. Jackson, M.: Problem Frames. Addison-Wesley, Reading (2001)
39. Russo, A., Nuseibeh, B., Kramer, J.: An abductive approach for analyzing event-based requirements specifications. In: Proc. of the 18th International Conference on Logic Programming (ICLP 2002), pp. 22–33. ACM, New York (2002)
40. Nuseibeh, B., Kramer, J., Finkelstein, A.: A framework for expressing the relationships between multiple views in requirements specifications. IEEE Transactions on Software Engineering 20(10), 760–773 (1994)

41. Nakatani, T., Fujino, T.: Role and owner based business domain analysis. In: Proc. of the Asia Pacific Software Engineering Conference 2005 (APSEC 2005), pp. 130–137. IEEE, Los Alamitos (2005)
42. Nakatani, T., Fujino, T.: Rodan: A requirements elicitation method for a business domain by focusing on roles. IPSJ Journal 48(8), 2534–2550 (2007) (in Japanese)
43. Brinkkemper, S., Saeki, M., Harmsen, F.: Meta-modelling based assembly techniques for situational method engineering. Information Systems 23(7), 489–508 (1998)
44. Checkland, P., Scholes, J.: Soft Systems Methodology in Action. John Wiley & Sons, Chichester (1991)
45. Yu, E.S.K., Mylopoulos, J.: An actor dependency model of organizational work –with application to business process reengineering. In: Proc. of the Conf. on Organizational Computing Systems (COOCS 1998), pp. 258–268 (1998)
46. Mizoguchi, R., Ikeda, M., Seta, K., Welkenhuysen, J.V.: Ontology for modeling the world from problem solving perspectives. In: Proc. of the International Joint Conferences on Artificial Intelligence 95 (IJCAI 1995) Workshop on Basic Ontological Issues in Knowledge Sharing, pp. 1–12 (1995)
47. Eriksson, H.E., Penker, M.: Business Modeling with UML. John Wiley & Sons, Chichester (2000)
48. Pree, W.: Design Patterns for Object-Oriented Software Development. Addison-Wesley, Reading (1996)
49. Haglind, M., Johansson, L., Rantzer, M.: Experiences integrating requirements engineering and business analysis. In: Proc. of the IEEE Third International Conference on Requirements Engineering, pp. 108–117. IEEE, Los Alamitos (1998)
50. Dardenne, A., van Lamsweerde, A., Fickas, S.: Goal-directed requirements acquisition. Science of Computer Programming 20, 3–50 (1993)
51. Jackson, M.: Seeing more of the world. IEEE Software 21(6), 83–85 (2004)
52. Carroll, J.M.: Making Use: Scenario-Based Design of Human-Computer Interactions. MIT Press, Cambridge (2000)
53. Alexander, I.: Misuse cases: Use cases with hostile intent. IEEE Software 20(1), 58–66 (2003)
54. Beus-Dukic, L., Myers, C.: Use and abuse cases. In: The 1st International Workshop on Requirements Engineering Education and Training, REET 2005 (2005), `http://re.cti.depaul.edu/REET05/F4_BeusDeukic.pdf`
55. Garfinkel, H., Rawls, A.: Ethnomethodology's Program. Rowman & Littlefield Publishers (2002)
56. CEDITI: A KAOS tutorial (2007), `http://www.objectiver.com/fileadmin/download/documents/KaosTutorial.pdf`
57. Object Management Group: Unified modeling language, `http://www.uml.org/`
58. Davis, A.M.: Just Enough Requirements Management: Where Software Development Meets Marketing. Dorset House, New York (2005)
59. Boehm, B., Bose, P., Horowitz, E., Lee, M.J.: Software requirements negotiation and renegotiation aides: A Theory-W based spiral approach. In: Proc. of the 17th International Conference on Software Engineering (ICSE 1995), pp. 243–253. ACM, New York (1995)
60. Phillips, J.J., Stone, R.D.: How to Measure Training Results – a Practical Guide to Tracking the Six Key Indicators. McGraw-Hill, New York (2002)

Chapter 7

Building Intelligent E-Learning Systems by Activity Monitoring and Analysis

Dumitru Dan Burdescu and Marian Cristian Mihăescu

University of Craiova, Romania

Abstract. E-Learning area has been intensively developed in recent years. One of the important research areas is related to improving e-Learning activity by giving the intelligent character to this activity besides core functionalities that is implemented in all e-Learning platforms.

This paper presents a method of providing intelligent character to an e-Learning platform by running a platform-side software module. The main goal of the module is to characterize learners according with performed activities and to offer advice regarding the resources that need to be accessed in order to increase the knowledge level of studied discipline. Acquiring this goal is accomplished by employing machine learning algorithms within platform-side software module. After learners are clustered based on performed activities, based on learner's activity parameters and parameters of target cluster there are obtained the resources which need more study. This approach is feasible due to the fact that the discipline is divided into chapters and each chapter has an associated concept map.

Keywords: E-Learning, activity monitoring and analysis, intelligent systems.

1 Introduction

There was designed and developed an e-Learning platform called Tesys (Burdescu, Mihaescu, 2006). This platform has implemented facilities for following type of users: system administrators, secretaries, professors and students. Some activities implemented for students, like downloading course materials or taking tests or exams are sometimes very heavy regarding the computational load of the server and the data traffic transfer to and from the user.

This paper presents the structure and functionality of an Expertise Module (EM) represented by a platform-side software module that runs along the Tesys e-Learning platform. The main purpose the EM is to provide the intelligent character for the educational network implemented by Tesys. The functionality of the EM module is presented in Figure 1.

As presented in Figure 1 the input of EM is represented by data traffic data. The data are obtained by a custom implemented logging mechanism embedded within the platform's business logic. The platform is represented by the setup put in place in order to perform all necessary activities within the e-Learning process. The setup

G.A. Tsihrintzis and L.C. Jain (Eds.): Multimedia Services in Intel. Environ., SIST 3, pp. 153–174.
springerlink.com

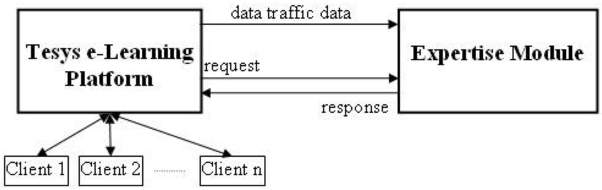


Fig. 1. General structure of intelligent e-Learning System

consists of course materials, test and exam quizzes that are set up by course managers and the overall setup performed by secretaries.

The Tesys e-Learning platform represents a collaborative environment in which all involved parties (e.g. secretaries, professors, students and administrators) accomplish their duties. The administrator, with the help of secretaries and professors are responsible for managing the environment in which the students will be through-out the e-Learning process. The platform has built in capability of monitoring and recording user's activity. The activity represents valuable data since it is the raw data for the machine learning and modeling process. The activity of each learner is seen as a sequence of sessions. A session starts when the student logs in and finishes when the student logs out. Under these circumstances, a sequence of actions makes up a session.

User's activity is monitored and recoded through a dedicated module implemented within the business logic of the platform. This facility was taken into consideration since the design phase of the platform. In was one of the requirements that the platform to be able to record user's performed actions with fine granularity.

The paper presents a platform-side software module that selects the resource(s) that need further attention of learner. Guiding the learning may have important benefits regarding the improvement of self-assessment effectiveness. The scope of the recommender system is making the learner obtain the maximum knowledge from the self-assessment activity. This is accomplished by the classifier according with all previous learners's performed activity. The activity is represented by the number of answers to questions regarding that concept, the average result of answered questions and the final result at the discipline. Each filtered resource is to be recommended or not. There will be defined the values each feature may have.

Following the structure of the discipline (chapters, concepts, and concepts maps) the professor creates a set of quizzes that may be accessed by the learner. Self-assessment activity is represented by taking a certain number of on-line quizzes. The scope of the recommender system is to guide the student to the resource he/she needs to access in order to make learning progress to be an effective one. The objective measure of accumulated knowledge is obtained from self-assessment activity. We think that this activity must be coordinated along with other learning activities. This coordination represents the means by which the recommender system makes the self-assessment as effective as possible.

The whole process is represented by an analysis procedure that has as input data representing the performed activities by learners. As learning algorithm it was employed Naive Bayes classifier (Mitchell, 1997). The classifier will predict the resources that the learner needs to access and study for improving his proficiency regarding the studied subject.

Regarding the design of software architecture of Tesys e-Learning platform many issues appear because it contains a mixture of data access code, business logic code, and presentation code. Such applications are difficult to maintain, because interdependencies between all of the components cause strong ripple effects whenever a change is made anywhere. The Model-View-Controller (Krasner, Pope, 1998) (MVC for short) design pattern solves these problems by decoupling data access, business logic, and data presentation and user interaction.

2 Related Work

A method of enhancing the quality of services offered by e-Learning platforms is the transformation of the environment into an adaptive one (Brusilovsky, 1996, 2001). An example of such system is AHA! (deBra, 2003). AHA! system implements a default adaptation technique by hiding less relevant hyper lines. Another research direction courses web-usage-based mining (Srivastava, 2000). In this direction there are major contributions regarding preprocessing and preparing of data (Cooley, 1999), recommendation of actions in e-Learning according performed actions (Zaiane, 2001, 2002).

There were proposed models for assisting evaluation of learner's in e-Learning systems (Guo, 2004). Implementing many of these research directions has been done using data mining algorithms (Abramovicz, 2004). There were employed clustering algorithms (Han and Camber, 2001), algorithms for obtaining sequential models (Srikant, Agrawal,1996), and algorithms for association rule creation (Tan, et. al. 2004; Agrawal, and Srikant, 1994). These research directions concretized into non invasive recommendation systems for learners (Lin, Alvarez, 2002; Spertus, Stein, 1998; Burdescu, Mihaescu, 2007). Such system is also employed for obtaining recommendations regarding the materials that need to be studied by learners (Tang, Mccalla, 2003). Implementation of such facilities has been accomplished by usage of dedicated pre-processing software fools that offer the possibility of automatization of performed operations (Marquardt, 2004).

Ausubel et. al. (1978) made the very important distinction between rote learning and meaningful learning. Meaningful learning requires three conditions: 1. The material to be learned must be conceptually clear and presented with language and examples relatable to the learner's prior knowledge. Concept maps can be helpful to meet this condition, both by identifying large general concepts held by the leaner prior to instruction of more specific concepts, and by assisting in the sequencing of learning tasks though progressively more explicit knowledge that can be anchored into developing conceptual frameworks; 2. The learner must possess relevant prior knowledge. This condition can be met after age 3 for virtually any domain of subject matter, but it is necessary to be careful and explicit in building concept frameworks if one hopes to present detailed specific knowledge in any field in subsequent lessons. We see, therefore, that conditions (1) and (2) are interrelated and both are important; 3. The learner must choose to learn meaningfully. The one condition over which the teacher or mentor has only indirect control is the motivation of students to choose to learn by attempting to incorporate new meanings into their prior knowledge, rather than simply memorizing concept definitions or propositional statements or computational procedures. The indirect control over this choice is primarily in instructional strategies used and the evaluation strategies used. Instructional strategies that emphasize relating new knowledge to the learner's existing knowledge foster meaningful learning. Evaluation strategies that encourage learners to relate ideas they possess with new ideas also encourage meaningful learning. Typical objective tests seldom require more than rote learning (Holden, 1992).

3 Concept Maps

The need for having well structured disciplines led to usage of concept maps. Concept maps are a result of Novak and Gowin's (Novak, Gowin, 1984) research into human learning and knowledge construction. Novak (Novak, 1977) proposed that the primary elements of knowledge are concepts and relationships between concepts are propositions. Novak (Novak, 1998) defined concepts as "perceived regularities in events or objects, or records of events or objects, designated by a label". Propositions consist of two or more concept labels connected by a linking relationship that forms a semantic unit. Concept maps are a graphical two-dimensional display of concepts (usually represented within boxes or circles), connected by directed arcs encoding brief relationships (linking phrases) between pairs of concepts forming propositions. The simplest concept map consists of two nodes connected by an arc representing a simple sentence such as 'flower is red,' but they can also become quite intricate.

One of the powerful uses of concept maps is not only as a learning tool but also as an evaluation tool, thus encouraging students to use meaningful-mode learning patterns (Mintzes, 2000; Novak, 1990; Novak, Gowin, 1984). Concept maps are also effective in identifying both valid and invalid ideas held by students, and this will be discussed further in another section. They can be as effective as more time-consuming clinical interviews for identifying the relevant knowledge a learner possesses before or after instruction (Edwards, Fraser, 1983).

Concept mapping may be used as a tool for understanding, collaborating, validating, and integrating curriculum content that is designed to develop specific competencies. Concept mapping, a tool originally developed to facilitate student learning by organizing key and supporting concepts into visual frameworks, can also facilitate communication among faculty and administrators about curricular structures, complex cognitive frameworks, and competency-based learning outcomes. To validate the relationships among the competencies articulated by specialized accrediting agencies, certification boards, and professional associations, faculty may find the concept mapping tool beneficial in illustrating relationships among, approaches to, and compliance with competencies (McDaniel, et. al. 1988).

Researchers suggest that learning can be enhanced by hands-on activities in which learners actively identify the attributes of concepts. For example, Harpaz, Balik, and Ehrenfeld (2004) recently demonstrated the efficacy of concept mapping in encouraging students to think independently and to find connections among concepts. Concept maps are powerful tools for helping learners identify general concepts prior to instruction focused on more specific content, as well as an evaluation tool of the knowledge that is formed afterward (Mintzes, et. al., 2000; Novak, Gowin, 1984).

In much the same way concept mapping has moved from being a knowledge representation to becoming a tool actively used by students during learning, curriculum developers have deployed concept maps during the initial curricular planning process. For example, Edmondson (1995) used concept mapping to develop a problem-based veterinary curriculum. He found that using concept mapping principles in the curricular planning process resulted in course content being more accessible and easily integrated by students. Concept mapping also provides educators a more comprehensive understanding of what students need to learn and helps eliminate sequencing errors in the development of lesson plans (Martin, 1994).

Recent decades have seen an increasing awareness that the adoption of refined procedures of evaluation contributes to the enhancement of the teaching/learning process. In the past, the teacher's evaluation of the pupil was expressed in the form of a final mark given on the basis of a scale of values determined both by the culture of the institution and by the subjective opinion of the examiner. This practice was rationalized by the idea that the principal function of school was selection - i.e. only the most fully equipped (outstanding) pupils were worthy of continuing their studies and going on to occupy the most important positions in society.

According to this approach, the responsibility for failure at school was to be attributed exclusively to the innate (and, therefore, unalterable) intellectual capacities of the pupil. The learning/ teaching process was, then, looked upon in a simplistic, linear way: the teacher transmits (and is the repository of) knowledge, while the learner is required to comply with the teacher and store the ideas being imparted (Vecchia, Pedroni, 2007).Usage of concept maps may be very useful for students when starting to learn about a subject. The concept map may bring valuable general overlook of the subject for the whole period of study. It may be advisable that at the very first meeting of students with the subject to include a concept map of the subject.

4 K-Means Clustering Algorithm

Clustering is one of the most useful tasks in data mining process for discovering groups and identifying interesting distributions and patterns in the underlying data. Clustering problem is about partitioning a given data set into groups (clusters) such that the data points in a cluster are more similar to each other than points in different clusters (Berry, M.J.A., Linoff, G., 1996).

In the clustering process, there are no predefined classes and no examples that would show what kind of desirable relations should be valid among the data that is why it is perceived as an unsupervised process (Berry, M.J.A., Linoff, G., 1996). On the other hand, classification is a procedure of assigning a data item to a predefined set of categories (Fayyad, M.U., et. al, 1996).

Clustering produces initial categories in which values of a data set are classified during the classification process.

From all clustering algorithms categories we chose to have a closer look on those that use partitioning methods. Firstly, k-Means algorithm is taken into consideration since is simple and straight forward. Still, a more realistic approach was needed to compensate the drawbacks of k-Means but still keeping the good parts. That is why fuzzy C-means algorithm was employed. The steps of clustering process are presented in Figure 2. The procedure follows the standard knowledge discovery (Fayyad, M.U., et. al, 1996) but is accustomed for our specific situation.

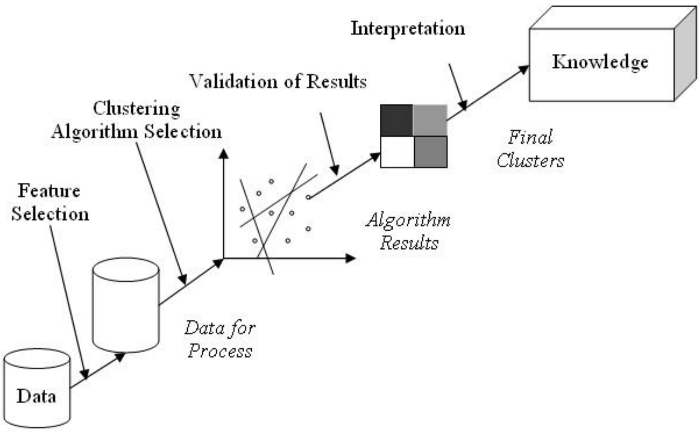


Fig. 2. Steps of clustering process

1) k-Means Algorithm: Given a database of n objects and k, the number of clusters to form, a partitioning algorithm organizes the objects into k partitions ($k \leq n$), where each partition represents a cluster. The clusters are formed to optimize an objective partitioning criterion, often called similarity function, such as distance, so that objects within a cluster are "similar", whereas the objects of different clusters are "dissimilar" in terms of database attributes. So, the first step is to define a list of attributes that may be representative for modeling and characterizing student's activity.

The classic k-means algorithm is a very simple method of creating clusters. Firstly, it is specified how many clusters are being thought: this is the parameter k. Then k points are chosen at random as cluster centers. Instances are assigned to their closest cluster centre according to he ordinary Euclidean function. Next the centroid, or the mean, of all instances in each cluster is calculated – this is the "means" part. These centroids are taken to be the new centre values for their respective clusters. Finally, the whole process is repeated with the new cluster centers. Iteration continues until the same points are assigned to each cluster in consecutive rounds, at each point the cluster centers have stabilized and will remain the same thereafter (Witten, Eibe, 2000).

From a different perspective for a cluster there may be computed the following parameters:

$$\mu = \frac{x_1 + x_2 + ... + x_n}{n} \text{, the means}$$

$$\sigma = \frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + ... + (x_n - \mu)^2}{n-1} \text{, the standard deviation}$$

p, the probability.

The sum of all probabilities for all clusters is 1. If we know which of the distributions each instance came from, finding the parameters is easy. On the other hand, if the parameters are known finding the probabilities that a given instance comes from each distribution is easy. Given an instance x , the probability that it belongs to cluster A is:

$$\Pr[A \mid x] = \frac{\Pr[x \mid A] - \Pr[A]}{\Pr[x]} = \frac{f(x;\mu_A,\sigma_A)p_A}{\Pr[x]}$$

Where $f(x;\mu_A,\sigma_A)$ is the normal distribution function for cluster A, that is:

$$f(x;\mu_A,\sigma_A) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

The EM algorithm takes into consideration that we know neither of these things: not the distribution that each training instance came from, nor the parameters μ, σ or the probability. So, we adopt the procedure used for the k-means clustering algorithm and iterate. Start with initial guess for the five parameters, use them to calculate the cluster probabilities for each instance, use these probabilities to estimate the parameters, and repeat. This is called the EM algorithm for "expectation-maximization". The first step, the calculation of cluster probabilities (which are the "expected" class values) is "expectation"; the second, calculation of the distribution parameters is "maximization" of the likelihood of the distributions given the data (Witten, Eibe, 2000).

Although the EM algorithm is guaranteed to converge to a maximum, this is a local maximum and may not necessarily be the same as the global maximum. For a better chance of obtaining the global maximum, the whole procedure should be repeated several times, with different initial guess for the parameter values. The overall log-likelihood figure can be used to compare the different final configuration obtained: just choose the largest of the local maxima (Witten, Eibe, 2000).

5 Software Architecture of Tesys E-Learning Platform

This architecture of the platform allows development of the e-learning application using MVC architecture. This three-tier model makes the software development process a little more complicated but the advantages of having a web application that produces web pages in a dynamic manner is a worthy accomplishment. The model is represented by DBMS (Data Base Management System) that in our case is represented by MySQL. The controller, which represents the business logic of the platform is Java based, being build around Java Servlet Technology. As servlet container Apache Tomcat 5.0 is used.

As for as we know there has not been yet solved the problem of development and integration of a software toolkit that uses data mining algorithms for analyzing actions performed by learners and which contributes to enhancement of quality of offered services by the e-Learning platform in which it has been integrated. This work presents a novel solution as well as generic for enhancement of quality of services implemented by e-Learning platforms. Researches within the project will be finalized by a software product that may be integrated into e-Learning platforms. The effort and resources involved in this research work of designing a toolkit that may be integrated into e-Learning platforms are continuous, efficient and adaptive to user's needs.

In the figure 3 there are presented the main software components from the MVC point of view.

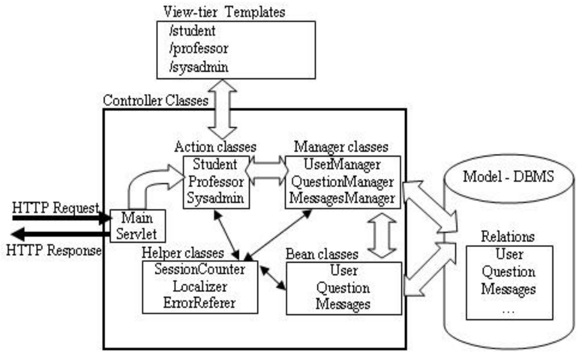


Fig. 3. Software components of the application from MVC point of view

MainServlet, Action, Manager, Bean, Helper and all Java classes represent the Controller. The Model is represented by the DBMS itself while the Webmacro templates represent the View. The model is built without any knowledge about views and controllers. We present platform from developer point of view.

The business logic of the application uses Java classes. As it can be seen in figure 1, there are four levels of dependency between classes. The levels are: servlets, actions, managers and beans.

The MainServlet can be seen as the main gate of the application. Its main purpose is to check whether the request may be fulfilled or not from user's role point of view. By default, a user is a visitor and may execute only non-critical actions. The MainServlet redirects visitor users to welcome page where they are supposed to provide a username and a password in order to login. All other requests are divided into three categories: login, logout and actions. When running a login or logout request the MainServlet delegates control to Login or Logout action classes. If the user is logged (i.e. he has provided a valid username and password) the MainServlet identifies his role within the application and for each requested it will check if he may run that action. If it can, MainServlet will delegate control to appropriate action class that will effectively fulfill the request.

The second level of classes in the dependency tree has the action classes that effectively take care of the requests. The main action classes are Login, Logout, Sysadmin, Secretary, Student and User.

The third level of classes in dependency tree has the manager classes that are extensively used by action classes. Manager classes are in close relationship with the database relations. We have implemented twelve manager classes for the most important relations. A manager class implements all needed operations on a specific relation. For example, QuestionsManager class has methods for insertion, deletion and updates a question. There are also methods that check whether a question exists or not, compute and save the result of a test or exam, save the questions that were part of the test or exam in database for later use. Finally, there are methods used to take a question from the database and prepare it for display or vice versa. The same type of methods may be found in manager classes for sections, disciplines, users, results, year structure, messages, and logs or blocked IPs.

The last level of classes consists of bean classes. These classes are used mainly in manager classes and are in one to one correspondence with the relations from the database. In fact, for each relation there was defined a bean class. We may say that an instance of a bean class represents a record from the correspondent relation from the database.

6 Monitoring Activity in Tesys E-Learning Platform

Tracking of actions is required to facilitate the data collection for the needed performed actions. The design of actions tracking module needs to select proper features that describe the performed actions of the learned. Among these features (or attributes) there are: user id, the date and time the action has been performed, details about the performed action, and so on. From the design phase of the platform, there were adopted two methodologies for monitoring actions.

Since the business logic of the platform is Java based, *log4j* utility package was employed as a logging facility and is called whenever needed within the logic of the application. The utility package is easy to use; *log4j.properties* properties file manages the logging process. The setup process states the logs are saved in *idd.log* file. The main drawback of this technique is that the data from the file is in a semi-structured form. This makes the information retrieval to be not so easy task to accomplish. On the advantages, logging activity may be very helpful in auditing the platform or even finding security breaches. This logging facility is also very helpful when debugging during development or when analyzing peculiar behavior during deployment.

To overcome the semi-structured shape of logged activity a structured way of gathering activity information was enforced. The activity table was added in the database and all actions were recorded in the manner of one record per action. In the table 1 it is presented the structure of activity table.

Table 1. Structure of activity table

Field	Description
id	primary key
userid	identifies the user who performed the action
date	stores the date when the action was performed
action	stores a tag that identifies the action
details	stores details about performed action
level	specifies the importance of the action

In Table 1 the action field is represented by a tag. The detailed explanation of what the tag means is set in a properties file. For each language a separate properties file is created, each file containing the same tags but with description in a different language.

The details field stores specific information regarding the action that was executed. For example, if a secretary modifies the profile of a student in the details field there will be stored information about what fields were updated. The level field specifies the importance of the executed action. There are defined three level of importance: 0, 1 and 2 where level 0 specifies the critical actions. After five months of deployment, the activity table contains more than 50,000 records and we suppose that until the end of the learning cycle there will be close to 100,000 records. All this logged activity may also be very helpful in an audit process of the platform. The records from the activity table represent the raw data of our analyzing process.

7 Data Filtering and Obtaining Recommendations

The decision model helps to determine whether a given resource is appropriate to be displayed to the learner as a recommendation for study. This is accomplished by clustering learners. The learners are clustered according with performed activity (e.g. number of taken tests), results (e.g. average of obtained grades), regarding resources (e.g. concepts, test quizzes, documents) that belong to a studied discipline.

The information about the resource and the user's activity are presented to the classifier as input, having as output a set of clusters of learners. Each learner belongs to only one cluster that gathers learners with similar performed activities and results. Then, it is identified the cluster with better performances and there are determined the resources that need improvement regarding performed activity and obtained results.

The training and testing data that are used for building the clusters is represented by instances. Each instance represents the activity of a learner regarding a resource. The structure of an instance is:

instance(studId, resourceId, noOfTests, avgTests)

The *studId* parameter identifies the learner who performed the activity. The *resourceId* parameter represents the resource for which the other features are valid. This is a numeric field. The *noOfTests* parameter represents the number of tests that were submitted by the learner regarding the resourceId resource. The *avgTests* parameter represents the average grade of tests that were submitted by the learner from the *resourceId* resource.

This is one of the setups used for experiments. Many other setups were used, using different sets of features (Mihaescu, Burdescu, 2006, 2007). In these work there were studied intensely how attribute selection influences results. Since in many practical situations there are far too many attributes for learning schemes to handle, and some of them – perhaps the overwhelming majority – are clearly irrelevant or redundant. Consequently, the data must be preprocessed to select a subset of attributes to use in learning. Of course, learning schemes themselves try to select attributes appropriately and ignore irrelevant and redundant ones, but in practice their performance can frequently be improved by preselection. For example, experiments show that adding useless attributes causes the performance of learning schemes such as decision trees and rules, linear regression, instance-based learners, and clustering methods to deteriorate (Witten, Eibe, 2000).

Obtaining recommendations is the last step of the analysis process. Once the model has been obtained and its quality has been acknowledged the effective process of obtaining recommendations may run. The acknowledgement of the quality of model is done by specifying thresholds regarding the percentage of correctly classified instances. These values are set up in a configuration file of the analysis module. In our setup, the model system stared making recommendations after the percentage of correctly classified instances raised above 80%.

Obtaining recommendations for students is accomplished by finding the resources that have values of parameters less than the ones of the computed target cluster. The target cluster is represented by other cluster than the one to which the learner belongs to and which has higher average values regarding number of taken tests and obtained grades. The current status of the student is evaluated at the moment he logs in the LMS. Then, the model selects the resources may be improved to higher values regarding the number of taken tests and obtained grades. Each time the learner accesses the LMS the current performed activity is matched against the current model and the corresponding resources whose parameters may be improved are displayed. After each week of running, the new data is considered for being included in training and testing data. This new data is added and new data and a challenger model is obtained. If the challenger model has better accuracy regarding the percentage of correct classified instance, it will replace the old model.

Sometimes it happens that during the usage of the system, the student could change his/her type of interaction with LMS. The problem of changes of the users' preferences is known as concept drift and has been discussed in several works about the use of machine learning for user modeling (Koychev, Schwab, 2000; Webb et. al., 2001). Concept drift can occur either because the acquired learning information needs to be adjusted or because the student simply changes his/her habits. In these scenarios, adaptive decision models, capable of better fitting the current student's preferences, are desirable.

8 Analysis Process and Experimental Results

The experimental results were obtained on Tesys e-Learning platform (Burdescu, Mihaescu, 2006). On this platform there was set Algorithms and Data Structures discipline. The tests were performed for five chapters: Simply/Double Linked Lists, Binary Search Trees, Height Balanced Trees, B Trees and Graphs. The first step was to build the General Concept Map by joining the concept maps from all chapters.

The concept map for Binary Search Trees is presented in figure 2. It contains 16 concepts, 11 linking phrases and 14 propositions.

The concepts are presented in table 2.

Table 2. List of Concepts

ID	Concept	ID	Concept
C1	BST	C9	Right child
C2	Dynamic Structure	C10	No child
C3	Node(s)	C11	Root
C4	Traversed	C12	Leaf
C5	Key	C13	Preorder
C6	Parent	C14	Inorder
C7	Child	C15	Postorder
C8	Left child	C16	Ascending order

The list of propositions with two concepts and one linking phrase is presented in table 3. The list of propositions with three concepts and two linking phrases is presented in table 4.

Table 3. List of propositions with two concepts and one linking phrase

Id	Concept	Linking phrase	Concept
P1	BST	is	Dynamic Structure
P2	BST	is made of	Node(s)
P3	Node	has	key
P4	Node	is	Parent
P5	Node	is	Child
P6	Parent	is greater than	Left child
P7	Parent	is smaller than	Right child
P8	Node	may have	Left child
P9	Node	may have	Right child
P10	Node	may have	No child

Table 4. List of propositions with three concepts and two linking phrase

Id	C	LP	C	LP	C
P11	Node	without	parent	is	Root
P12	BST	may be	traversed	in	Preorder
P13	BST	may be	traversed	in	Inorder
P14	BST	may be	traversed	in	Postorder

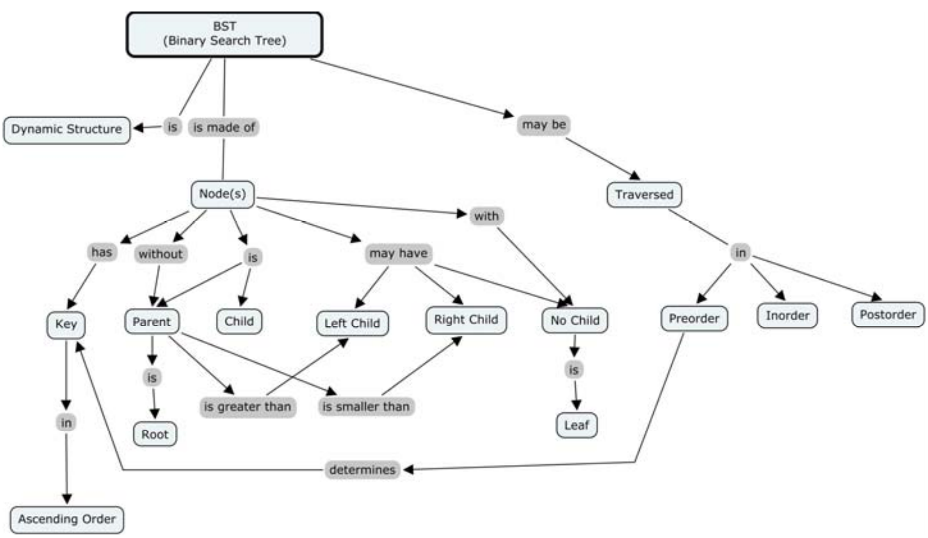


Fig. 4. Binary Search Tree Concept Map

Once the concept map has been built the general graph of the each chapter may be created. In this graph, each proposition will become an edge that links the first concept and the last concept. The domain knowledge expert will assign a weight for each edge. While the students answers questions the number of correct and wrong answers will determine the knowledge weight of that edge.

There is one proposition with five concepts and four linking phrases: “BST” may be “Traversed” in “Preorder” determines “Key” in “Ascending Order”. The concepts are bolded and put between quotation marks, while linking phrases are italic and underlined.

Knowledge evaluation is closely related with cognitive processes performed by an individual. After an initial step of goal setting a student has at first to identify task-relevant knowledge and to evaluate it with respect to his own knowledge regarding that goal. Self-evaluation of individual knowledge is a step that should be performed before any learning process. For example, if the task is to acquire expert knowledge, the structure of an individuals' knowledge as represented in an individual knowledge map may be compared with the knowledge structure of an expert as represented in an expert map. The potential of knowledge maps as means for diagnosing individual structures of knowledge has been shown in a variety of empirical studies (Jonassen et. al., 1997). In self-regulated learning scenarios the particular contribution of computer-based concept maps is that they may support self-assessment (Kommers, Lanzing, 1997; Shavelson et. al., 1994).

A concept map may be seen as an oriented graph where vertexes are represented by concepts and edges are represented by verbs. Within e-Learning platform for each proposition from the concept map may will be represented by a weighted edge and will have associated a number of quiz questions. Under these circumstances we have created an algorithm for building the associated graph of a concept map. The parameters of edges are continuously updated as the student answers quizzes. In the experimental part of the paper there will be presented the obtained graph. Each edge in the graph will have four parameters: the weight, the total number of questions, the correctly answered questions and the wrong answered questions.

Knowledge evaluation procedure takes into account the parameters of edges from the associated graph of concept map. The weight of an edge is set by the domain knowledge expert from a scale from 1 to 10 where 1 means very simple proposition and 10 means very hard proposition. All other parameters take different values according with learner’s experience. In the experimental part there will be presented the formulas that synthesize the knowledge level of the learner.

The analysis of concept’s map associated graph represents the core part of decision support system that runs along the e-Learning platform. The architecture of the decision support system is presented in figure 5.

Each chapter has associated a concept map build by the domain expert. From the concept map a transformation procedure creates the general graph of the chapter. In this graph, each sentence becomes an edge, weighted by the domain expert. Besides the associated weight, each proposition has associated a set of quiz questions that are to be answered by learners.

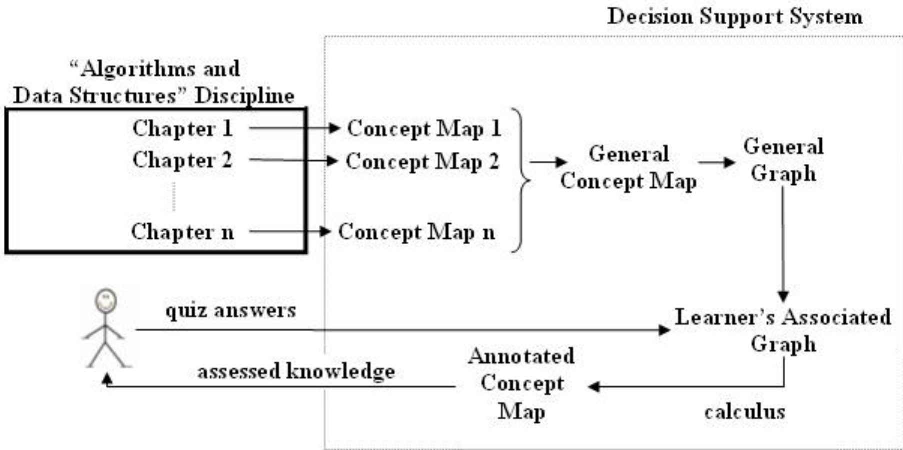


Fig. 5. Functionality of Decision Support System

When the learner starts answering questions, the Decision Support System starts building learner's associated graph. This graph represents the input data for the calculus procedure that assesses the knowledge of the students regarding each chapter and the discipline as a whole. Whenever a student logs in the Decision Support System builds the learner's associated graph such that at request the knowledge status will be delivered in the form of an annotated concept map. The query of the annotated map returns information regarding the concept coverage at chapter and at discipline level. It is the option of the domain expert to specify different weights to chapters. This situation occurs when the domain expert considers that concepts from one chapter are more important than concepts covered by other chapter. Under these circumstances for each chapter there will be assigned a certain percentage such that the sum of percentages for all chapters is 100.

If:

W- is the weight of the edge

CA – is the number of correct answers

WA – is the number of wrong answers

N – the number of questions

Than

KW – is the knowledge weight of the edge

and

$$KW = \frac{CA - WA}{N} \frac{1}{W} * 100$$

Under these circumstances the knowledge weight may also be negative. At any time there may be estimated the overall knowledge level of the learner as the ratio between overall knowledge weight and overall weight. Figure 4 presents the general graph associated with the concept map.

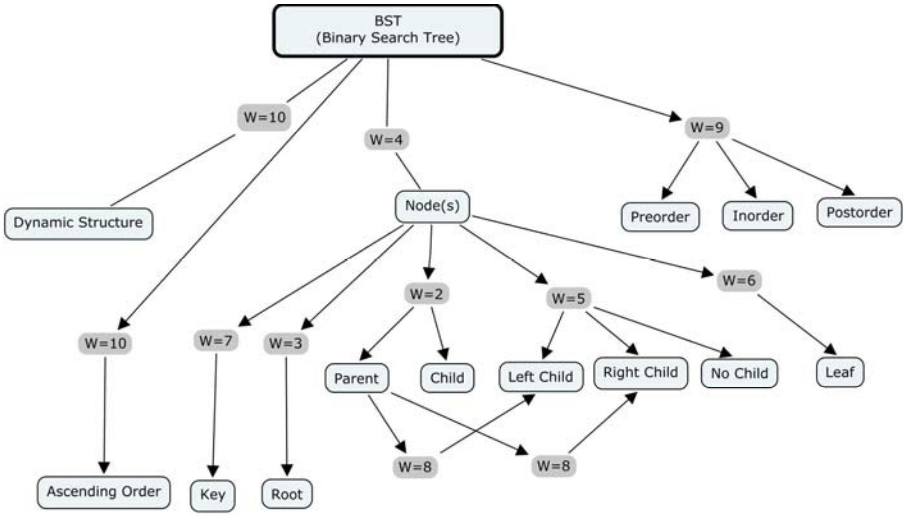


Fig. 6. Binary Search Tree General Graph

The algorithm transforming the Concept Map into General Graph is strait forward. Each proposition becomes an edge with a weight assigned by domain knowledge expert. In this way it was obtained the Binary Search Tree General Graph.

Once the General Graph has been set up the professor has to set up the quiz questions for the chapter. For each edge in the graph it will correspond a certain number of quiz questions. There is no specification regarding the number of quiz questions but a minimum (e.g. five) number is still required. Once the quiz questions have been set up, for each student there may be constructed the learner's associated graph. This graph will have associated with the edges the history of correct and wrong answered questions. The Calculus engine will reconstruct an Annotated Concept Map which will present to the learner the current status of his knowledge level at Concept level. In this way, the learner will have an exact overview of his knowledge level regarding that chapter.

The Annotated Concept Map may represent the important information for learner in having a decision regarding which part of the chapter needs more study.

Table 5 presents a sample of the setup of the Binary Search Trees chapter.

Table 5. Sample setup of BST chapter

Proposition	Weight	No. of questions
P1	10	8
P2	4	7
P3	7	6
P4	3	5
P5	2	7

Table 6. Sample values for Learner's Associated Graph

Proposition (Weight)	**No. of questions**	**CA**	**WA**	**KW (%)**
P1 (10)	8	3	2	1.25
P2 (4)	7	4	2	7.14
P3 (7)	6	1	3	-4.76
P4 (3)	5	3	1	13.3
P5 (2)	7	2	4	-14.2

Table 6 presents sample values of the Learner's Associated Graph corresponding to BST chapter.

The values from table 5 are marked in an Annotated Concept Map that is finally presented to the learner. The Annotated Concept Map is the final outcome of the Decision Support System and is supposed to guide the learner regarding the necessary future efforts.

Table 7 presents the weights of chapters as they were assigned by the domain expert.

Table 7. Sample Weights Assigned to Chapters

Chapter	**Weight**
Simply/Double Linked Lists	15
Binary Search Trees	15
Height Balanced Trees	25
B Trees	25
Graphs	20

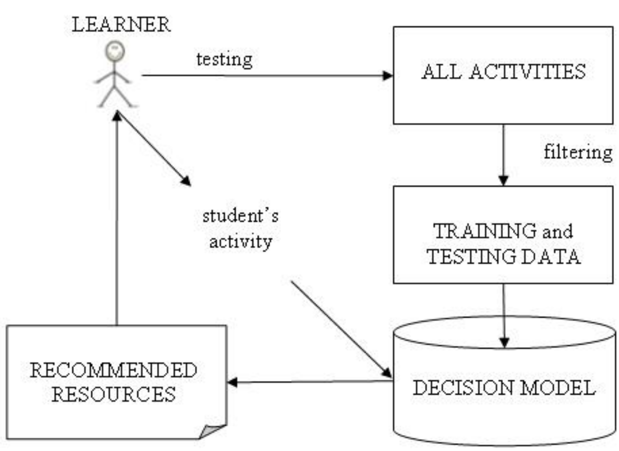


Fig. 7. The selection of recommended resources

The analysis process for building the model is presented in figure 5.

Firstly, it is solved the cold-start problem. This means the data is collected such that the decision model is created. All performed activities are filtered such that training and testing data is obtained. Because we use Weka the data is extracted and translated into a standard format called ARFF, for Attribute Relation File Format (Garner et. al. 1995; Holmes et. al., 1994). This involves taking the physical log files and database relations and processing them through a series of steps to generate an ARFF dataset.

The values of attributes are computed for each instance through a custom developed off-line Java application. The outcome of running the application is in the form of a file called *activity.arff* that will later be used as source file for Weka workbench (Weka, 2008).

The *activity.arff* file has a standard format which is composed of two sections. In the first one there is defined the name of the relation and the attributes. For each attribute there is defined the set of nominal values it may have. In the next lines it is presented the first section of the file.

@relation activity

@attribute *studId* INTEGER

@attribute *resourceId* INTEGER

@attribute *noOfTests* {1, 2, 3, 4, 5}

@attribute *avgTests* {1, 2, 3, 4, 5}

In this section of the file are defined all attributes. An important decision that is needed is to establish the granularity for each attribute which is represented by the number of nominal values it may take. As it can be seen from the above presented lines we consider five possible values for *noOfTests* feature. In the same there are defined the set of possible values for each of the features.

The second section of the *activity.arff* file is represented by the data itself. Here are all the instances that will enter the classification process. In the next lines there are presented few instances that may be found in this section.

instance(studId, resourceId, noOfTests, avgTests)

@data

1, 1, 2, 3

1, 2, 3, 2

2, 3, 4, 3

Each row represents an instance. For example, the first row represents an instance (learner with id equal to 1) which used resource 1 for taking very many tests and obtained high grades for tests regarding this. In the same way there can be interpreted all other instances. At this point we may say we have obtained useful data that may be used for experimentation with machine learning schemes.

The K-Means algorithm was run in Weka. The original dataset was divided into a training of 90% of instances and a test set of 10 % of instances. The model was constructed using four attributes: *studId, resourceId, noOfTests, avgTests.*

More detailed results regarding the obtained model are presented below.

```
=== Run information ===
Scheme:       weka.clusterers.EM -I 100 -N 3 -S 100 -M 1.0E-6
Relation:     activity
Instances:    928 (268 students)
Attributes:   4
              studId, resourceId, noOfTests, avgTests
Test mode:    evaluate on training data
=== Model and evaluation on training set ===
EM
==
Number of clusters: 3
Clustered Instances (learners)
0     91 ( 34%)
1     42 ( 16%)
2     135 ( 50%)
Log likelihood: -2.61092
```

The first cluster contains 91 instances (343%), the second cluster contains 42 instances (16%) and the third cluster contains 135 instances (50%).

The performance of the model was evaluated computing the logarithm of similarity value for which it was obtained a value of -2.61092. This value is small enough to represent a good and also promising result.

9 Discusson and Conclusion

In this chapter we have presented an analysis process that clusters students according with their performed activity. The recommender system determines a cluster with students that had better performances and determines the resources that need more study. The clustering has been used in Tesys e-Learning system for obtaining the resources of a discipline the learner needs to study. To discover the user's needs there is used information about all previous learner's performed activities. The advantages of using these information is that this allows refining the initial beliefs acquired by the system by observing the student's performed actions over time thus computing up-to-date learning activity for each student. For classification of resources there was used k-Means clustering algorithm as the machine learning algorithm for determining to which cluster a learner belongs. The experiments carried out to obtain an initial model are described in detail.

Tesys e-Learning platform has been designed such that on-line testing activities may me performed as they were set up by course managers. It has been created a Concept Map for a Binary Search Trees chapter as well as for each chapter of Algorithms and Data Structures course. The Concept maps have been the staring point in creating the sets of quiz questions. Each quiz question refers to a certain proposition from the concept map.

For the designed Concept Maps it has been derived a General Graph in which edges are represented by the propositions from the General Concept Map. For each edge the domain knowledge expert (i.e. course manager) assigned a specific weight. A weight has been assigned for each chapter such that the sum of all weights is 100.

We have created a procedure of data analysis which may provide interesting conclusions regarding the classification of students from an e-learning platform.

The platform was developed, is currently running and has built in capabilities of monitoring students activities. An off-line application was developed for creating the input data files that are analyzed. Data analysis is done using EM clustering algorithm implemented by Weka system. The main goal is characterizing of students on one hand and the platform on the hand.

Student's characterization may have a predictive value in the sense that from the activities a student has made there may be pulled conclusions about his learning proficiency. On the other hand, platform's characterization may have as result an estimation of the capability of an e-learning system to grade and order students according to their accumulated knowledge. This analysis is critical for having as conclusion that a system can support generalized tests.

After the setup has been put in place, the learners started using the platform. At request, from the general graph there was derived the learner's associated graph and on this one there may be performed calculus such that the level of knowledge regarding the resource may be estimated. These calculus represent the annotations in the original concept. The annotated concept map represents what the learner finally receives upon his request.

This whole mechanism represents the functionality of a decision support system that runs along the Tesys e-Learning platform. For each learner there is initialized a decision model from data obtained from all previous learner's performed activities. The outcome of the Decision Support System is represented by a set of recommended resources that are presented to the learner.

Each learner's individual decision model is then adapted from the observations of the learner's performed activities. Moreover, the model is also able to adapt itself to changes in the learner's performed activities. Knowledge was obtained by employing an automated discovery process of iterative manner. A dataset was taken and analyzed using the facilities provided by Weka workbench for manipulating datasets, configuring learning schemes and handling output.

For learners, it was observed that students that followed the recommendations had better results in shorter time due to optimization of their learning scenarios. After new more data is collected a challenger decision model may be created and replace the old one if its accuracy is higher.

Within Tesys e-Learning platform set up there were registered 268 students. After six month of running the platform there were over 40.000 recorded actions that represent the raw data. Filtering this data had as output an arff file with 928 instances.

To summarize, the results from our experiments show the feasibility of machine learning in improving self-assessment effectiveness by activity monitoring and analysis. The machine learning approach gives a wide field of action in the area of adaptive e-Learning platforms due to the wide variety of models that may be used for description or prediction. For learners, the main benefit regards improvement in their meta-cognitive skills to observe, analyze and improve their learning process. For course managers, this approach helps understanding and interpreting the ongoing course at individual learner level and also at group level.

In future there may be tested different setups, with more test questions and adjusted weights. The weights may be adjusted by comparing the end results with the estimated results. In this way, there may be discovered resources that need increase or decrease in weight.

As future works there are two directions that may be discussed. One regards the used features and the other regards the employed algorithms. Since the range of activities is much wider than the one used in this chapter, there may be performed new experiments with different sets of features. For example, features like the number of sent messages to course manager may be added to the feature set.

References

Abramovicz, W., Kaczmarek, T., Kowalkiewicz, M.: Supporting topic map creation using data mining techniques. Australian Journal of Information Systems (2004)

Agrawal, R., Srikant, R.: Fast Algorithms for Mining Association Rules. In: Proc. of the 20th Int. Conf. on Very Large Databases, Santiago, Chile (1994)

ATutor (2008), http://atutor.com

Ausubel, D.P., Novak, J.D., Hanesian, H.: Educational psychology: A cognitive view, 2nd edn. Holt, Rinehart and Winston, New York (1978)

Bennett, P.N.: Assessing the calibration of Naive Bayes' posterior estimates. Technical Report No. CMU-CS00-155 (2000)

Berry, M.J.A., Linoff, G.: Data Mining Techniques For Marketing, Sales and Customer Support. John Wiley & Sons, Inc., USA (1996)

Blackboard (2008), http://www.blackboard.com

Brusilovsky, P.: Methods and techniques of adaptive hypermedia. User Modeling and User Addapted Interaction (1996)

Brusilovsky, P.: Adaptive Hypermedia. User Modeling and User Adapted Interaction (2001)

Burdescu, D.D., Mihăescu, M.C.: Tesys: e-Learning Application Built on a Web Platform. In: Proceedings of International Joint Conference on e-Business and Telecommunications, Setubal, Portugal, pp. 315–318 (2006)

Burdescu, D.D., Mihăescu, M.C.: Enhancing the Assessment Environment within a Learning Management Systems. In: EUROCON 2007 – The International Conference on "Computer as a Tool", Warsaw, Poland, pp. 2438–2443 (2007)

Cooley, R., Mobasher, B., Srivastava, J.: Data preparation for mining World Wide Web browsing patterns. Knowledge and Information Systems (1999)

de Bra, P., Berden, B., de Lange, B., Rousseau, B., Santic, T., Smits, D., Stash, N.: AHA! The Adaptive Hypermedia Architecture. In: ACM Conference on Hypertext and Hypermedia, Nottingham, UK (2003)

Edmondson, K.M.: Concept Mapping for the development of medical curricula. Journal of Research in Science Teaching 32(7), 777–793 (1995)

Edwards, J., Fraser, K.: Concept maps as reflections of conceptual understanding. Research in Science Education 13, 19–26 (1983)

Fayyad, M.U., Piatesky-Shapiro, G., Smuth, P., Uthurusamy, R.: Advances in Knowledge Discovery and Data Mining. AAAI Press, Menlo Park (1996)

Frank, E., Trigg, L., Holmes, G., Witten, I.H.: Naive Bayes for Regression. Machine Learning 41(1), 5–15 (2000)

Garner, S.R., Cunningham, S.J., Holmes, G., Nevill-Manning, C.G., Witten, I.H.: Applying a Machine Learning Workbench: Experience with Agricultural Databases. In: Proc Machine Learning in Practice Workshop, Machine Learning Conference, Tahoe City, CA, USA, pp. 14–21 (1995)

Guo, L., Xiang, X., Shi, Y.: Use web usage mining to assist background online e-Learning assessment. In: 4th IEEE ICALT (2004)

Han, J., Camber, M.: Data Mining Concepts and Techniques. Morgan Kaufman, San Francisco (2001)

Harpaz, I., Balik, C., Ehrenfeld, M.: Concept Mapping: An educational strategy for advancing nursing education. Nursing Forum 39(2), 27–30 (2004)

Holden, C.: Study flunks science and math tests. Science Education 26, 541 (1992)

Holmes, G., Donkin, A., Witten, I.H.: Weka: a machine learning workbench. In: Proceedings of the 1994 Second Australian and New Zealand Conference on Intelligent Information Systems, Brisbane, Australia, pp. 357–361 (1994)

Jonassen, D.H., Reeves, T.C., Hong, N., Harvey, D., Peters, K.: Concept mapping as cognitive learning and assessment tools. Journal of Interactive Learning Research 8(3/4), 289–308 (1997)

Kommers, P., Lanzing, J.: Student's concept mapping for hypermedia design. Navigation through the world wide web (WWW) space and self-assessment. Journal of Interactive Learning Research 8(3/4), 421–455 (1997)

Kononenko, I.: Comparison of Inductive and Naive Bayesian Learning Approaches to Automatic Knowledge Acquisition. In: Current Trends in Knowledge Acquisition. IOS Press, Amsterdam (1990)

Koychev, I., Schwab, I.: Adaptation to Drifting User's Interests. In: Proceedings of ECML 2000 Workshop: Machine Learning in New Information Age (2000)

Krasner, G.E., Pope, S.T.: A cookbook for using the model-view-controller user interface paradigm in smalltalk-80. JOOP (1998)

Lin, W., Alvarez, S.: Efficient adaptive-support association rule mining for recommender systems. Data Mining and Knowledge Discovery (2002)

Marquardt, C.G.: A preprocessing tool for web usage mining in distance education domain. In: IDEAS: Proc. of the 8th Int. Database Engineering and Applications Symposium (2004)

Martin, D.J.: Concept Mapping as an aid to lesson planning: A longitudinal study. Journal of Elementary Science Education 6(2), 11–30 (1994)

McDaniel, E., Roth, B., Miller, M.: Concept Mapping as a Tool for Curriculum Design. Issues in Informing Science and Information Technology (1988)

Mihaescu, M.C., Burdescu, D.D.: Testing Attribute Selection Algorithms for Classification Performance on Real Data. In: 3rd IEEE International Conference on Intelligent Systems (IS), pp. 581–586. IEEE Press, Los Alamitos (2006)

Mihaescu, M.C., Burdescu, D.D.: Classification of students using their data traffic within an e-Learning platform. In: International Joint Conference on e-Business and Telecommunications -International Conference on e-Business (ICE-B), pp. 315–321. INSTICC Press (2007)

Mintzes, J.J., Wandersee, J.H., Novak, J.D.: Assessing science understanding: A human constructivist view. Academic Press, San Diego (2000)

Mitchell, T.: Machine Learning. McGraw-Hill, New York (1997)

Moodle (2008), `http://www.moodle.com`

Novak, J.D.: A Theory of Education. Cornell University Press, Ithaca (1977)

Novak, J.D., Gowin, D.B.: Learning How to Learn. Cambridge University Press, New York (1984)

Novak, J.D.: Concept maps and vee diagrams: Two metacognitive tools for science and mathematics education. Instructional Science 19, 29–52 (1990)

Novak, J.D.: Learning, Creating, and Using Knowledge: Concept Maps as Facilitative Tools in Schools and Corporations. Lawrence Erlbaum Associates, Mahwah (1998)

Shavelson, R.J., Lang, H., Lewin, B.: On concept maps as potential "authentic" assessments in science (Technical Report 388). UCLA, Center for the Study of Evaluation (CSE/CRESST), Los Angeles (1994)

Srikant, R., Agrawal, R.: Mining sequential patterns: Generalizations and performance improvements. In: EDBT: Proc. of teh 5th Int. Conf. On Extending Database Technologies (1996)

Spertus, E., Stein, L.A.: A hyperlink-based recommender system written in squeal. In: Proc. ACM CIKM 1998 Workshop on Web Information and Data Management (1998)

Srivastava, J., Cooley, R., Deshpande, M., Tan, P.N.: Web usage mining: Discovery and applications of usage patterns from web data. SIGKDD Explorations (2000)

Tan, P.N., Kumar, V., Srivastava, J.: Selecting the right objective measure for association analysis, Information Systems (2004)

Tang, T., Mccalla, G.: Smart recommendation for evolving e-Learning system. In: Proc. Workshop on Technologies for Electronic Documents for Supporting Learning, Int. Conf. on Artificial Intelligence in Education, Sydney, Australia (2003)

Vecchia, L., Pedroni, M.: Concept Maps as a Learning Assessment Tool. Issues in Informing Science and Information Technology (2007)

Witten, I.H., Eibe, F.: Data Mining – Practical Machine Learning Tools and Techniques with Java Implementations. Morgan Kaufmann Publishers, San Francisco (2000)

Webb, G., Pazzani, M., Billsus, D.: Machine Learning for User Modeling. User Modeling and User-Adapted Interaction, 19–29 (2001)

WebCT (2008), `http://www.webct.com`

Weka (2008), `http://www.cs.waikato.ac.nz/ml/weka`

Zaiane, O.R.: Web usage mining for a better web-based learning environment. In: Proc. of. Conf. on Advanced Technology for Education, Banff, Alberta (2001)

Zaiane, O.R.: Building a recommender agent for e-Learning systems. In: Proc. of 7th Int. Conf. on Computers in Education, Auckland, New Zeeland (2002)

Chapter 8

Supporting the Search for Explanations of Medical Exceptions

Rainer Schmidt[1] and Olga Vorobieva[1,2]

[1] Institute for Medical Informatics and Biometry, University of Rostock, Germany
rainer.schmidt@medizin.uni-rostock.de
[2] Sechenov Institute of Evolutionary Physiology and Biochemistry, St.Petersburg, Russia

Abstract. Doctors are confronted with many exceptions. In their practice they have to consider them and have to deal with them appropriately. In research, often exceptions occur which contradict a hypothesis. Statisticians are happy as long as an appropriate test shows a necessary significance, whereas doctors are unhappy about such a situation, especially when the number of exceptional cases is rather high.

We have developed two systems that give support for both of these situations. The first one deals with therapy inefficacy. In medical practice, therapies that are prescribed according to a certain diagnosis sometimes do not give desired results. Sometimes therapies are effective for some time but suddenly stop helping any more. There are many different reasons. For long-term therapy support in the endocrine domain and in psychiatry, we have developed a Case-Based Reasoning system that not only performs typical therapeutic tasks but also especially deals with situations where therapies become ineffective.

The second systems helps to explain cases that do not fit into a theoretical hypothesis. It deals with situations where neither a well-developed theory nor reliable knowledge nor a proper case base is available. Instead, just some theoretical hypothesis and a set of measurements is given. Usually a few cases do not satisfy the research hypothesis. We examine all these exceptional cases to find explanations. The system offers a dialogue to guide the search for possible reasons in all components of the data system. In this chapter, an application on "dialyse patients" is presented, who where offered a specially developed complex of physiotherapy exercises. The purpose of this fitness offer was to improve the physical conditions of the patients and to increase the quality of their lives. The hypothesis that fitness improves the condition of dialyse patients was statistically confirmed. However, the conditions of a couple of patients worsened though they participated in the fitness program. Our system was used by a medical expert to find out why.

1 Introduction

In their medical practice doctors are very often confronted with exceptional cases. Systems that are intended to support doctors have to take such situations into account. In knowledge-based systems such exceptions have to be considered and dealt with

G.A. Tsihrintzis and L.C. Jain (Eds.): Multimedia Services in Intel. Environ., SIST 3, pp. 175–199.
springerlink.com

appropriately. In rule-based systems this often leads to a huge amount of exceptional rules, namely very every possible exception an appropriate specific rule.

However, exceptions do not only occur in medical practice but also in medical research. In medical studies, statisticians are usually quite happy when an appropriate statistical test significantly supports a hypothesis. Especially when the number of cases contradicting a hypothesis is rather big, physicians are often not satisfied with this situation. They often additionally wish to get explanations for the exceptional cases.

In this chapter, we present two systems that deal with both of these situations. The first system is called ISOR-1. It deals with all sorts of therapeutic tasks and it especially investigates therapy inefficacy. This is done within a conversational dialogue with the user, where user and system can inspire each other.

The second system is called ISOR-2. It helps to explain cases that contradict a theoretical research hypothesis. This is also done within a conversational dialog of the user and the system.

Both systems make use of the Case-Based Reasoning method. That means they considered former already solved cases (exceptions) when attempting to find explanation for an exceptional query case. In this chapter, both systems are presented separately, one after the other, first ISOR-1, subsequently ISOR-2.

2 ISOR-1: Investigating Therapy Inefficacy

ISOR-1 deals with long-term diseases, especially with some psychiatric diseases and with endocrine malfunctions. For specific psychiatric diseases, Case-Based Reasoning systems have already been developed, namely for Alzheimer's disease [1] and for eating disorders [2].

Though ISOR-1 can already cope with a couple of problems, in this chapter we mainly focus on just two endocrine and psychiatric disorders, namely on hypothyroidism and depressive symptoms. For hypothyroidism just one therapy is available, namely Levothyroxine. The difficulty is to compute an appropriate dose.

Inefficacy of pharmacological therapy for depression is a widely known problem [3, 4]. There are many approaches to solve this problem. Guidelines and algorithms have been created [5, 6]. ISOR gives reference to a psychopharmacology algorithm [6] that is available on the website htp://mhc.com/Algorithms/Depression.

Apart from computing initial therapies and doses there two more therapeutic tasks, namely later on a dose update if necessary, and the consideration of interactions with further diseases, complications, and especially with already administered.

To compute initial doses for hypothyroidism (figure 1), there exist a couple of guidelines [7]. The assignment of a patient to a fitting guideline is obvious because of the way these guidelines have been defined. By applying them a range for good doses can be calculated. To compute a dose with best expected impact we retrieve similar cases whose initial doses are within the calculated ranges. Since cases are described by few attributes and since our case-base is rather small, we use a sequential measure of dissimilarity, namely the one proposed by Tversky [8]. On the basis of those retrieved cases that had best therapy results an average initial therapy is calculated. Best therapy results can be determined by values of a blood test after two weeks of treatment with the initial dose.

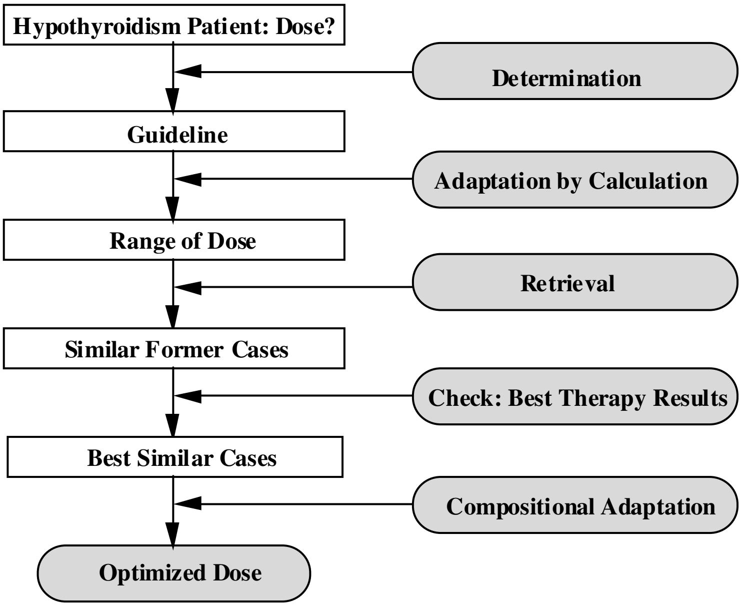


Fig. 1. Computation of an initial dose

Updating the dose in a patient's lifetime. For monitoring a hypothyroidism patient, three basic laboratory blood tests (TSH, FT3, and FT4) have to be undertaken. Usually, the results of these tests correspond to each other. Otherwise, it indicates a more complicated thyroid condition and additional tests are necessary. If the results of the basic tests show that the thyroid hormone level is normal, it means that the current levothyroxine dose is OK. If the tests indicate that the thyroid hormone level is too low, the current dose has to be increased by 25 or 50 µg, if it is too high, the dose has to be decreased by 25 or 50 µg [9]. So, for monitoring, adaptation means a calculation according to some rules, which are based on guidelines. Since an overdose of levothyroxine may cause serious complications for a patient, a doctor cannot simply consider test results and symptoms that indicate a dose increase but additionally has to investigate reasons why the current dose is not appropriate any more. In ISOR-1 this situation is described as a problem of therapy inefficiency.

Additional diseases or complications. It often occurs that patients do not only have hypothyroidism, but they suffer from further chronic diseases or complications. Thus, a levothyroxine therapy has to be checked for contraindications, adverse effects and interactions with additionally existing therapies (figure 2). Since no alternative is available to replace levothyroxine, if necessary additionally existing therapies have to be modified, substituted, or compensated [9].

ISOR-1 performs three tests. The first one checks if another existing therapy is contraindicated to hypothyroidism. This holds only for very few therapies, namely for specific diets like soybean infant formula, which is the most popular food for babies who do not get enough mother's milk but it prevents the effect of levothyroxine. Such

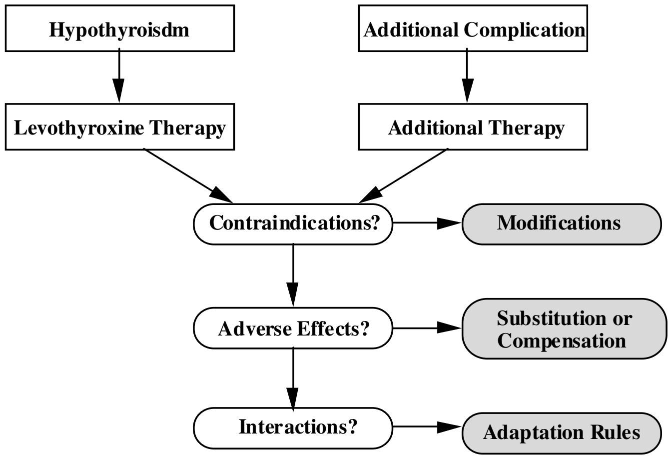


Fig. 2. Levothyroxine therapy and additional therapies

diets have to be modified. Since no exact knowledge is available to explain how to accomplish this, just a warning is issued, which says that a modification is necessary.

The second test considers adverse effects. There are two ways to deal with them. A further existing therapy has either to be substituted or it has to be compensated by another drug. Such knowledge is available, and we have implemented corresponding rules for substitutional and compensational adaptation.

The third test checks for interactions between both therapies. We have implemented some adaptation rules, which mainly attempt to avoid the interactions. For example, if a patient has heartburn problems that are treated with an antacid, a rule for this situation states that levothyroxine should be administered at least four hours after or before an antacid. However, if no adaptation rule can solve such an interaction problem, the same substitution rules as for adverse effects are applied.

Apart from these therapeutic tasks, the main task of ISOR-1 is the investigation of therapy inefficacy. In medical practice, therapies prescribed according to a certain diagnosis sometimes do not give desired results. Sometimes therapies are effective for some time but then suddenly stop helping any more. There are many possible reasons. A diagnosis might be erroneous, the state of a patient might have changed completely, or the state might have changed just slightly but with important implications for an existing therapy. Furthermore, a patient might have caught an additional disease, some other complication might have occurred, a patient might have changed his/her lifestyle (e.g. started a diet) and so on.

So, ISOR-1 is a medical Case-Based Reasoning system that deals with the following tasks:

- choose appropriate (initial) therapies,
- compute doses for chosen therapies,
- update dose recommendations according to laboratory test results,
- establish new doses of prescribed medicine according to changes in a patient's medical status or lifestyle,

- find out probable reasons why administered therapies are not as efficient as they should be,
- test obtained reasons for inefficacy and make sure that they are the real cause, and
- suggest recommendations to avoid inefficacy of prescribed therapies.

2.1 System Architecture

ISOR-1 is designed to solve typical therapy problems, especially inefficacy of prescribed therapies that can arise in many medical domains. Therefore most algorithms and functions are domain independent. Another goal is to cope with situations where important patient data are missing and/or where theoretical domain knowledge is controversial.

ISOR-1 does not generate solutions itself. Its task is to help users by providing all available information and to support them when they search for optimal solutions. Users shall be doctors, maybe even together with a patient.

In addition to the typical Case-Based Reasoning knowledge, namely former already solved cases, ISOR-1 uses further knowledge components, namely medical histories of query patients themselves and prototypical cases (prototypes). Furthermore, ISOR-1's knowledge base consists of therapies, conflicts, instructions etc. The idea of combining case-bases with different knowledge courses was, for example, already used in Creek [10]. ISOR-1's general program scheme is shown in figure 3.

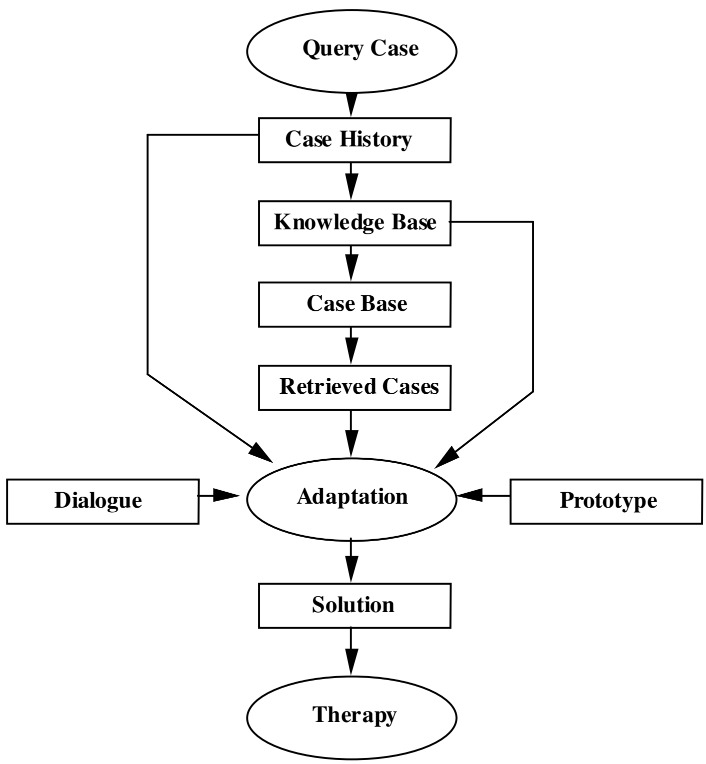


Fig. 3. General program scheme of ISOR-1

2.1.1 Medical Case Histories

Ma and Knight [11] have introduced a concept of case history in Case-Based Reasoning. Such an approach is very useful when we deal with chronic patients, because often the same complications occur again, former successful solutions can be helpful again, and former unsuccessful solutions should be avoided.

The case history is written in the patient's individual base in form of a sequence of records. A patient's base contains his/her whole medical history, all medical information that is available: diseases, complications, therapies, circumstances of his/her life, and so on. Each record describes an episode in a patient's medical history. Episodes often characterise a specific problem. Since the case-base is problem oriented, it contains just episodes and the same patient can be mentioned in the case-base a few times, even concerning different problems.

Information from the patient's individual base can be useful for a current situation, because for patients with chronic diseases very similar problems often occur again. If a similar situation is found in the patient's history, it is up to the user to decide whether to start retrieval in the general case-base or not.

In endocrinology, case histories are designed according to a standard scheme, one record per visit. Every record contains the results of laboratory tests and of an interrogatory about symptoms, complaints and physiological conditions of a patient. Therefore the retrieval of former similar situations from the individual base of an endocrine patient is easy to organise.

For psychiatric patients, case histories are often unsystematic and they can be structured in various forms. A general formalisation of psychiatric cases and their histories is not achieved yet. The design of case histories is problem dependent.

In both domains, we first search in the query patient's history for similar problems and for similar diagnoses.

2.1.2 Knowledge Base, Case-Base, and Prototypes

The knowledge base contains information about problems and their solutions that are possible according to the domain theory. It has a tree structure and it consists of lists of diagnoses, corresponding therapies, conflicts, instructions, and medical problems (including solutions) that can arise from specific therapies. The knowledge base also contains links to guidelines, algorithms and references to correspondent publications [5, 6].

The case-base is problem oriented. Thus a case in the case-base is just a part of a patient's history, namely an episode that describes a specific problem that usually has a solution too. So, the case-base represents decisions of doctors (diagnosis, therapies) for specific problems, and their generalisations and their theoretical foundations (see the examples in the following section). A solution is called "case solution". Every case solution has (usually two) generalisations, which are formulated by doctors. The first one is expressed in terms of the knowledge base and it is used as a keyword for searching in the knowledge base. Such a generalisation is called "knowledge base solution". The second generalisation of a solution is expressed in common words and it is mainly used for dialogues. It is called "prompt solution".

Former cases (attribute value pairs) in the case-base are indexed by keywords. Each case contains keywords that have been explicitly placed by an expert. For retrieval three main key words are used: a code of the problem, a diagnosis, and a therapy. Further key words such as age, sex and so on can be used optionally.

Prototypes (generalized cases) play a particular role. Prototypes help to select a proper solution from the list of probable or available solutions. A prototype may help to point out a reason of inefficacy of a therapy or it may support the doctor's choice of a drug.

2.1.3 Retrieval, Adaptation, and Dialogue

For retrieval keywords are used. Since ISOR-1 is problem oriented, the first keyword is a code that implies a specific problem. The second one is the diagnosis and the other ones are retrieved from the knowledge base.

Adaptation takes place as a dialogue between the doctor and the system. The idea of dialogue interaction with the user is similar to conversational CBR [12]. The system presents different solutions, versions of them, and asks questions to manifest them. The doctor answers and selects suggestions, while the patient himself suggests possible solutions that can be considered by the doctor and by ISOR-1.

We differentiate between two steps of adaptation. The first one occurs as a dialogue between ISOR-1 and the user. Usually, doctors are the users. However, sometimes even a patient may take part in this dialogue. The goal of these dialogues is to select probable solutions from all information sources mentioned in the sections above. Pieces of information are retrieved by the use of keywords. Specific menus support the retrieval process. The first step of adaptation can be regarded as partly user based, because ISOR-1 presents lists of probable solutions and menus of keywords, the user selects the most adequate ones. The second adaptation means proving obtained solutions. This proving is rule based and it includes further dialogues, laboratory test results, and consultations with medical experts. While the procedures supporting the first adaptation step are domain independent, the adaptation rules of the second step are mainly domain dependent.

2.2 Illustration

Now we illustrate by three examples how ISOR-1 works. The first and the second example are from the endocrine domain, the third one deals with a psychiatric problem.

2.2.1 Inefficacy of Levothyroxine Therapy

Every morning a mother gives her 10 year-old boy not only the prescribed Levothyroxine dose but also vitamin pills. These pills have not been prescribed but they are healthy and have lately been advertised on TV. Part of this medication is sodium hydrocarbonate (cooking soda) that causes problems with Levothyroxine.

<u>Problem:</u> Levothyroxine does not help any more.

<u>Individual base.</u> The same problem, inefficacy of Levothyroxine therapy, is retrieved from the patient's history. The solution of the former problem was that the boy did not take the drug regularly. This time it must be a different cause, because the mother controls the intake.

Knowledge base. It has a tree structure that is organised according to keys. One main key is *therapy* and the keyword is *Levothyroxine*. Another keyword is *instructions*. These instructions are represented in form of rules that concern the intake of Levothyroxine. For Levothyroxine a rather long list of instructions exists. Since the idea is that the boy may break an instruction rule, this list is sorted according to the observed frequency of offences against them in the case-base.

Concerning these instructions a couple of questions are asked, e.g. whether the boy takes sodium hydrocarbonate together with Levothyroxine. Since the mother is not aware of the fact that Sodium hydrocarbonate is contained in the vitamin pills, she gives a negative answer and no possible solution can be established by the knowledge base. However, *soda* is generated as one keyword for retrieval in the case-base.

Eventually, three general solutions are retrieved from the knowledge base: sodium hydrocarbonate, soy, and estrogen. Since estrogen does not fit for the boy, it is eliminated.

Case-base. Using the keyword *soda* eight cases with seven individual solutions are retrieved.

Thus we get a list of drugs and beverages that contain sodium hydrocarbonate. All of them belong to the generalised solution "soluble".

Solution. The boy admits to take Levothyroxine together with an instantiation of the generalised solution "soluble", namely soluble vitamin.

Recommendation. The boy is told to take vitamin four hours later than Levothyroxine. Additionally, further interactions between vitamin and Levothyroxine must be checked, because it might be necessary to adjust the Levothyroxine dose.

2.2.2 Improving the Efficacy by Dose Updates

Figure 4 shows an example of a case study. We compared the decisions of an experienced doctor with the recommendations of ISOR-1. The decisions are based on basic laboratory tests and on lists of observed symptoms. Intervals between two visits are approximately six months. In this example, there are three deviations between the doctor's and ISOR-1's decisions, usually there are less. At the second visit (v2), according to laboratory results the Levothyroxine should be increased. ISOR-1 recommended a too high increase. The applied adaptation rule was not precise enough. So, we modified it. At visit 10 (v10) the doctor decided to decrease the dose. The doctor's reasons were not included in our knowledge base and since his attempt was not successful, we did not alter any adaptation rule.

At visit 21 (v21) the doctor increased the dose because of some minor symptoms of hypothyroidism, which were not included in ISOR-1's list of hypothyroidism symptoms. Since the doctor's decision was probably right (visit 22), we added these symptoms to the list of hypothyroidism symptoms of ISOR-1.

2.2.3 Inefficacy of Psychiatric Drugs

Originally, ISOR-1 was developed for the endocrine domain, especially for hypothyroidism, but later on it has been generalised. Now it can solve the same types of problems in different medical domains. The following example comes from the psychiatry domain.

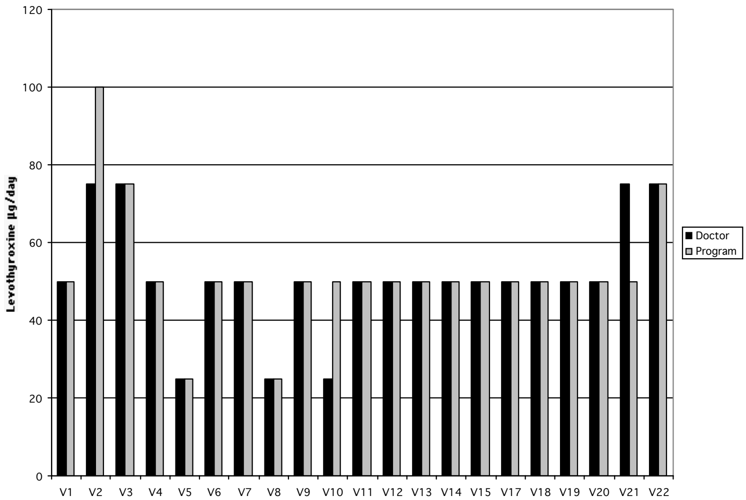


Fig. 4. Dose updates recommended by our program compared with doctor's decision. V1 means the first visit, V2 the second visit etc.

A man, 55 years of age, has been treated for depression for 15 years. Symptoms of depression appeared after he was cruelly beaten near his house. Since he did not see any connection between his depression and the violence, he did not tell it to his doctor. At first, the antidepressant Tofranil for intake in the morning and the sedative Tisercin for better sleep at bedtime were prescribed. Later on, during another depression stage the antidepressant Fluoxetine helped. Now, his problem is that neither Fluoxetine, nor any other proposed antidepressant helps any more.

Retrieval. Keywords are used to retrieve information from all data sources. Since optional keywords about a patient's feeling (e.g. *feeling worse*) are used for retrieval in the patient's medical history, even biographical events of a patient can be retrieved.

Individual base. Since the problem of inefficacy of an antidepressant never arose in the patient's past, no solution can be found. However, indirect information was retrieved. The keyword *feeling better* provided a trip to Switzerland, while the result of the keyword *feeling worse* provided a trip to Israel, where the latest very severe depression began.

Feeling better: A trip to Switzerland
Feeling worse: A trip to Israel

The knowledge base contains information about depression, anxiety and other psychiatric diseases, possible complications and references to their theoretical grounds [13, 14]. References to similar problems are retrieved, the most remarkable one is a link to the algorithm for psychopharmacology of depression [6]. Though the idea of the algorithm is to solve the problem of non-response to an antidepressant, it does not really fit here, because it does not cover the situation that a therapy helped for some time and then stopped having an effect.

Case-base. Eleven cases with similar depression problems are retrieved. Three of them are characterised by the general idea *depression is secondary to anxiety resulting from a psychical trauma*. These three cases are:

Case solution 1: A severe stress during World War 2 (a woman)
Case solution 2: A bad experience in jail (a young man)
Case solution 3: Sexual assault in childhood (a woman)

All other cases have solutions that are generalised to *changes in therapy*.

Adaptation. ISOR-1 displays retrieved information pieces. In this case, two strategies are offered. The first one suggests trying some other therapy. This strategy is supported by the majority of the retrieved cases and partly by theoretical recommendations. The second strategy means to check the diagnosis. This strategy is supported by the three retrieved cases mentioned above and by the patient's medical history. The choice between both strategies is up to the user. In this example, the doctor chooses to attempt the second strategy at first, because various therapy changes have already been attempted in the past – without success. Furthermore, the doctor is especially led by the patient's medical history, because Switzerland is usually associated with a safe life (especially in comparison to life in Russia), while living in Israel is considered as unsafe. Additionally, this strategy is supported by the general situation that some sedative drugs (Neuroleptics) had helped for some time (e.g. Tisercin at the beginning).

ISOR-1 offers a list of questions for the favoured strategy and as a result the doctor concludes that in this case depression is in fact only second to anxiety. The man is permanently afraid of possible violence and anxiety is based on strong fear that occurred long ago.

Explaining remarks. Diagnosing anxiety needs good medical skills, because patients try to suppress traumatic events from their memory [15]. In this example, depression even served as a mechanism of suppression. The accepted case-based solution spared the patient unnecessary experiments with other psychopharmacological drugs.

So, the first problem is solved, a new diagnosis is ascertained.

The next problem is prescription of a therapy. According to the domain theory and to our knowledge base anxiety implies Neuroleptics [14, 16]. Many of them are available but a good choice is not trivial.

Individual base. From the patient's history those sedatives (Neuroleptics) are retrieved that he took in his lifetime and that had positive effects on his psychical condition: Tisercin and Paxil, which is a drug that has both sedative and antidepressive effects (figure 5).

Prototype. Among those prototypes that have been defined by doctors (based on their long experience with cases) the prototypical solution Paxil is retrieved.

Adaptation. Before described, every drug must be checked for conflicts with the patient's additional diseases and already existing therapy. Though the query patient has already taken Paxil in the past, ISOR-1 checks all possible conflicts. If necessary, adaptation has to be performed. In this case no conflicts are discovered and Paxil is prescribed.

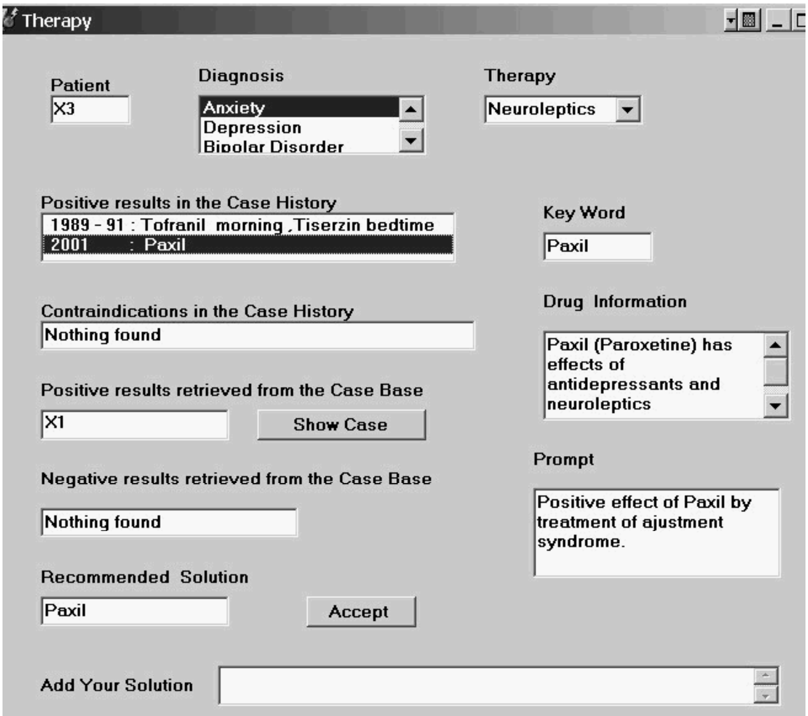


Fig. 5. Example of the therapy recommendation menu

3 ISOR-2: Explaining Exceptional Cases

ISOR-2 is a logical continuation of ISOR-1. It is still the same system and the same structure of dialogues, but ISOR-2 deals with situations where neither a well-developed theory nor reliable knowledge nor a proper case-base is available. So, instead of reliable theoretical knowledge and intelligent experience, just a theoretical hypothesis and a set of measurements are given. In such situations the usual question is, how do measured data fit to a theoretical hypothesis. To statistically confirm a hypothesis it is necessary, that the majority of cases fit the hypothesis. Mathematical statistics determines the exact quantity of necessary confirmation [17]. However, usually a few cases do not fit the hypothesis. We examine these exceptional cases to find out why they do not satisfy the hypothesis. ISOR-2 offers a dialogue to guide the search for possible reasons in all components of the data system.

This approach is justified by a certain mistrust of statistical models by the doctors, because modelling results are usually unspecific and “average oriented” [18], which means a lack of attention to individual "imperceptible" features of concrete patients.

The usual Case-Based Reasoning assumption is that a case-base with complete solutions is available. Our approach starts in a situation where such a case-base is not available at the beginning but has to be set up incrementally (figure 6). So, we must

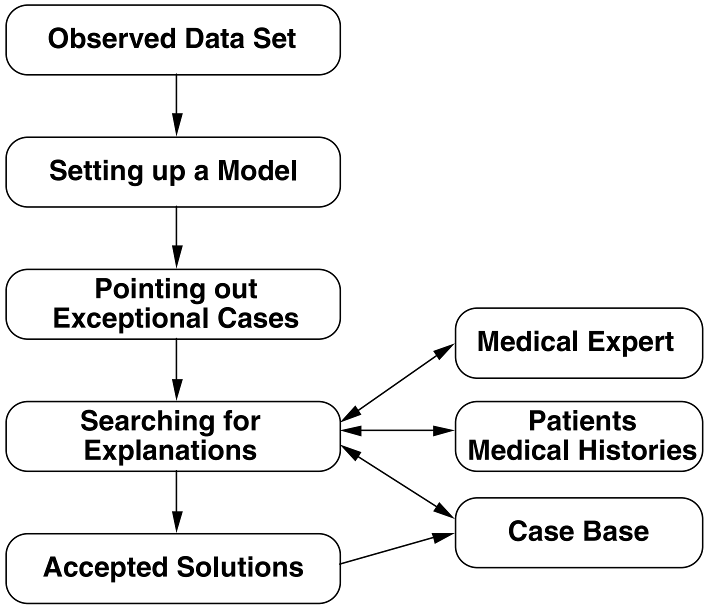


Fig. 6. The general program scheme of ISOR-2

1. Construct a model,
2. Point out the exceptions,
3. Find causes why the exceptional cases do not fit the model, and
4. Set up a case-base.

Case-Based Reasoning is combined with a model, in this specific situation with a statistical one. The idea to combine CBR with other methods is not new. Care-Partner, for example, resorts to a multi-modal reasoning framework for the co-operation of CBR and Rule-based Reasoning (RBR) [19]. Another way of combining hybrid rule bases with CBR is discussed by Prentzas and Hatzilgeroudis [20] (2002). The combination of CBR and model-based reasoning is discussed in [21]. Statistical methods are used within CBR mainly for retrieval and retention [22, 23]. Arshadi proposes a method that combines CBR with statistical methods like clustering and logistic regression [24].

3.1 Dialyse and Fitness

Hemodialyse means stress for a patient's organism and has significant adverse effects. Fitness is the most available and relative cheap way of support. It is meant to improve the physiological condition of a patient and to compensate negative dialyse effects.

One of the intended goals of this research is to convince the patients of the positive effects of fitness and to encourage them to make efforts and to go in for sports actively. This is important because dialyse patients usually feel sick, they are physically weak, and they do not want any additional physical load [25].

At the University clinic in St. Petersburg, a specially developed complex of physiotherapy exercises including simulators, walking, swimming etc. was offered to all dialyse patients but only some of them actively participated, whereas some others participated but were not really active.

3.2 Data

For each patient a set of physiological parameters is measured. These parameters contain information about burned calories, maximal power achieved by the patient, his oxygen uptake, his oxygen pulse (volume of oxygen consumption per heartbeat), lung ventilation and others. There are also biochemical parameters like haemoglobin and other laboratory measurements. More than 100 parameters were measured for every patient.

Parameters are supposed to be measured four times during the first year of participating in the fitness program. There is an initial measurement followed by a next one after three months, then after six months and finally after a year. Unfortunately, since some measurements did not happen, many data are missing. Therefore the records of the patients often contain different sets of measured parameters.

It is necessary to note that parameter values of dialyse patients essentially differ from those of non-dialysis patients, especially of healthy people, because dialyse interferes with the natural, physiological processes in an organism. In fact, for dialyse patients all physiological processes behave abnormal. Therefore, the correlation between parameters differs too.

For statistics, this means difficulties in applying statistical methods based on correlation and it limits the usage of a knowledge base developed for normal people. Inhomogeneity of observed data, many missing data, and many parameters for a relatively small sample size, all this makes the data set practically impossible for usual statistical analysis.

Since the data set is incomplete, additional or substitutional information has to be found in other available data sources. These are databases – the already existent individual base, the sequentially set up case-base, and the medical expert as a special source of information.

3.3 Setting up a Model

We start with a medical problem that has to be solved based on given data. In this application it is: "Does special fitness improve the physiological condition of dialyse patients?" More formal, physical conditions of active and non-active patients have to be compared. The patients are divided into two groups, depending on their activity, active patients and non-active ones.

According to our assumption active patients should feel better after some months of fitness, whereas non-active ones should feel rather worse. The meaning of “feeling better” and “feeling worse” has to be defined in this context. A medical expert selects appropriate factors from ISOR’s menu. The expert selected the following main factors:

- F1: O2PT - oxygen pulse by training
- F2: MUO2T - maximal uptake of oxygen by training
- F3: WorkJ – performed work (joules) during control training

Subsequently the “research time period” has to be determined. Initially, this period was planed to be twelve months, but after a while some patients tend to give up the fitness program. This means, the longer the time period, the more data are missing.

Therefore, we have to make a compromise between time period and sample size. A period of six months has been chosen.

The next question is whether the model shall be quantitative or qualitative? The observed data are mostly quantitative measurements. The selected factors are of quantitative nature too. On the other side, the goal of this research is to find out whether physical training improves or worsens the physical condition of the dialyse patients.

It is not necessary to compare one patient with another patient. Instead, each patient has to be compared with his own situation some months ago, namely just before the start of the fitness program. The success shall not be measured in absolute values, because the health statuses of patients are very different. Thus, even a modest improvement for one patient may be as important as a great improvement of another. Therefore, the development is simply classified in two categories: “better” and “worse”. Since the usual tendency for dialyse patients is to worsen in time, those few patients, where no changes could be observed, are added to the category ”better”.

The three main factors are supposed to describe the changes of the physical conditions of the patients. The changes are assessed depending on the number of improved factors:

- Weak version of the model: at least one factor has improved
- Medium version of the model: at least two factors have improved
- Strong version of the model: all three factors have improved

The final step means to define the type of model. Popular statistical programs offer a large variety of statistical models. Some of them deal with categorical data. The easiest model is a 2x2 frequency table. The “better/ worse” concept fits this simple model very well. The results are presented in table 1.

Table 1. Results of Fisher’s Exact Test, performed with an interactive Web-program: http://www.matforsk.noIola/fisher.htm

Improvement mode	Patient’s physical condition	Active	Non-active	Fisher Exact p
Strong	Better	28	2	< 0.0001
	Worse	**22**	21	
Medium	Better	40	10	< 0.005
	Worse	**10**	12	
Weak	Better	47	16	< 0.02
	Worse	**3**	6	

Unfortunately, the most popular Pearson Chi-square test is not applicable here because of the small values “2” and “3” in table 1. But Fisher’s exact test [17] can be used. In the three versions shown in table 1, a very strong significance can be observed. The smaller the value of p is, the more significant the dependency.

So, the performed Fisher test confirms the hypothesis that patients doing active fitness achieve better physical conditions than non-active ones. However, there are exceptions, namely active patients whose health conditions did not improve.

Exceptions should be explained. Incrementally, explained exceptions build the case-base. According to table 1, the stronger the model, the more exceptions can be observed and have to be explained. Every exception is associated with at least two problems. The first one is “Why did the patient’s condition get worse?” Of course, “worse” is meant in terms of the chosen model. Since there may be some factors that are not included in the model but have changed positively, the second problem is “What has improved in the patient’s condition?” To solve this problem, significant factors, where the values improved, have to be found.

In the following section we explain the set-up of a case-base on the strongest model version.

3.4 Setting Up a Case-Base

We intend to solve both problems (mentioned above) by means of CBR. So, a case-base has to be set up sequentially. That means, as soon as an exception is explained, it is incorporated into the case-base and can help to explain further exceptional cases. A random order is chosen for the exceptional cases. In fact, they are taken in alphabetical order.

The retrieval of already explained cases is performed by keywords. The main keywords are the usual ISOR ones, namely “problem code”, “diagnosis”, and “therapy”. Here, the instantiations of these keywords are “adverse effects of dialysis” (diagnosis), “fitness” (therapy), and two specific problem codes.

Besides the main ISOR-2 keywords additional problem specific ones are used. Here, the additional keyword is the number of worsened factors. Further keywords are optional. They are just used when the case-base becomes bigger and retrieval is not simple any longer.

However, ISOR-2 does not only use the case-base as knowledge source but further sources are involved, namely the patient’s individual base (his medical history) and observed data (partly gained by dialogue with medical experts). Since in the domain of kidney disease and dialyse the medical knowledge is very detailed and much investigated but still incomplete, it is unreasonable to attempt to create an adequate knowledge base. Therefore, a medical expert, observed data, and just a few rules serve as medical knowledge sources.

3.4.1 Expert Knowledge and Artificial Cases

Expert’s knowledge can be used in many different ways. It can be used to acquire rules, to select appropriate items from the list of retrieved solutions, to propose new solutions, and last but not least – to create artificial cases.

Initially, an expert creates artificial cases, afterwards they can be used in the same way as real cases. They are created in the following situation. An expert points out a factor F as a possible solution for a query patient. Since many values are missing, it can happen that just for the query patient values of factor F are missing. In this case, the doctor's knowledge can not be applied. However, it is sensible to save it anyway. Principally, there are two different ways to do this. The first one means to generate a correspondent rule and to insert it into ISOR-2's algorithms. Unfortunately, this is very complicated, especially to find an appropriate way for inserting such a rule. The alternative is to create an artificial case. Instead of a patient's name an artificial case number is generated. The other attributes are either inherited from the query case or declared as missing. The retrieval attributes are inherited within the dialog (see the fields eleven and twelve of figure 7). ISOR-2's algorithms remain intact. Artificial cases can be treated in the same way as real cases, they can be revised, deleted, generalised and so on.

3.4.2 Solving the Problem "Why Did Some Patients Conditions Became Worse?"

As results we obtain a set of solutions of different origin and different nature. There are three solution categories: additional factor, model failure, and wrong data.

Additional factor. The most important and most frequent solution is the influence of an additional factor. Only three main factors are obviously not enough to describe all cases. Unfortunately, for different patients different additional factors are important. When ISOR-2 has discovered an additional factor as explanation for an exceptional case, the factor has to be confirmed by the expert before it can be accepted as a solution. One of these factors is Parathyroid Hormone (PTH). An increased PTH level sometimes can explain a worsened condition of a patient [6]. PTH is a significant factor, but unfortunately it was measured only for some patients.

Another additional factor solution is a phosphorus blood level. We use the principle of artificial cases to introduce the factor phosphorus as a new solution. One patient's record contains many missing data. The retrieved solution means high PTH, but PTH data in the current patient's record is missing too. The expert proposes an increased phosphorus level as a possible solution. Since data about phosphorus data are also missing, an artificial case is created, who inherits all retrieval attributes of the query case whereas the other attributes are recorded as missing. According to the expert, high phosphorus can explain the solution. Therefore it is accepted as an artificial solution or a solution of an artificial case.

Model failure. We regard two types of model failures. One of them is deliberately neglected data. As a compromise we just consider six months as research period but later data of a patient might be important. In fact, three of the patients did not show an improvement in the considered six months period but in the following six months. So, they are wrongly classified and shall really belong to the "better" category.

The second type of model failure is based on the fact that the two-category model is not precise enough. Some exceptions can be explained by a tiny and not really significant change in one of the main factors.

Wrong data are usually due to a technical mistake or to data that has not really been proved. One patient, for example, was reported as actively participating in the fitness program but really was not.

3.4.3 Solving the problem "What in a Patient's Condition Became Better?"

There are at least two criteria to select factors for the model. First, a factor has to be significant, and second there must be enough patients for which this factor was measured at least for six months. So, some principally important factors are initially not taken into account because of missing data.

The list of solutions includes these factors: haemoglobin, maximal power (watt) achieved during control training. Oxygen pulse and oxygen uptake were measured in tow different situations, namely during the training under loading and before training in a state of relax. Therefore we have two pairs of factors: oxygen pulse in state of relax (O2PR) and during training (O2PT); maximal oxygen uptake in state of relax (MUO2R) and during training (MUO2T). Measurements made in a state of relax are more indicative and significant than those made during training. Unfortunately, most measurements were made during training. Only for some patients correspondent measurements in relax state exist. Therefore O2PT and MUO2T are accepted as main factors and are taken into the model. On the other side, O2PR and MUO2R serve as solutions for the current problem.

In the case-base every patient is represented by a set of cases, every case represents a specific problem. This means that a patient is described from different points of view and accordingly different problem keywords are used for retrieval.

3.5 Illustration of ISOR-2's Program Flow

Figure 7 shows the main dialogue of ISOR-2. For illustration purposes, we have numbered the fields of the menu. At first, the user sets up a model (fields one to four), subsequently he/she gets the result (field 5) and an analysis of the model (fields six to eight), and then he/she attempts to find explanations for the exceptions (fields nine and ten). Finally, the case-base is updated (fields eleven and twelve). Now the steps are explained in detail.

At first, the user has to set up a model. To do this, he has to select a grouping variable. In the example of figure 7, CODACT is chosen. It stands for "activity code" and means that active and none active patients are to be compared. Provided alternatives are the sex and the beginning with the fitness program (within the first year of dialyse or later). In another menu, the user can define further alternatives. Furthermore, the user has to select a model type (alternatives are "strong", "medium", and "weak"), the length of the research period (3, 6, or 12 months), and the main factors have to be selected. The list contains all factors from the observed database. In the example of figure 7, three factors are chosen: O2PT (oxygen pulse by training), MUO2T (maximal oxygen uptake by training), and WorkJ (work in joules during the test training). In the menu list, the first two factors have alternatives: "R" instead of "T", where "R" stands for state of rest and "T" stands for state of training.

When the user has selected these items, the program calculates the table. "Better" and "worse" are meant in the sense of the chosen model, in the example the strong model has been chosen. ISOR-2 does not only calculate the table but additionally extracts the exceptional patients from the observed database. In the menu, the list of exceptions shows the code names of the patients. In the example of figure 7, patient "D5" is selected" and all further data belong to this patient. The goal is to find an explanation for this exceptional case "D5". In field seven of the menu it is shown that

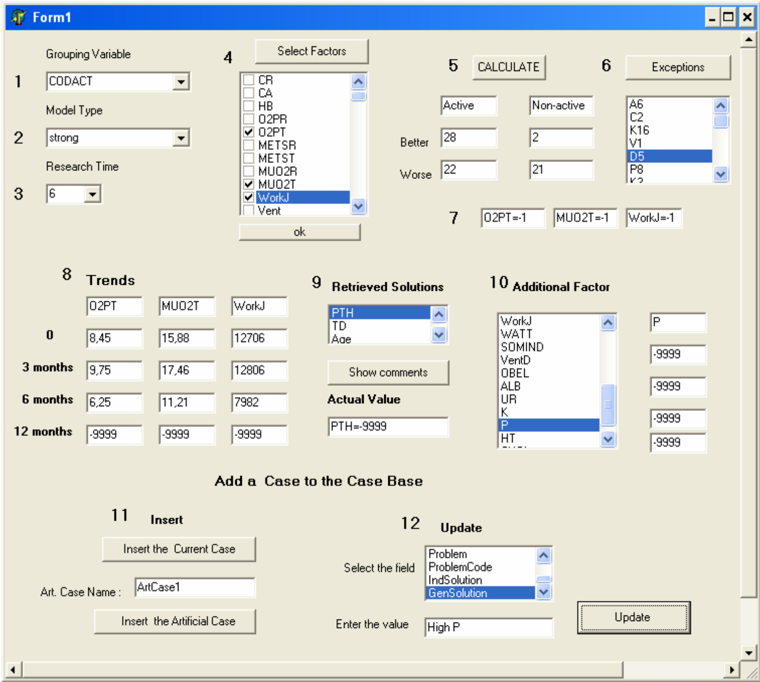


Fig. 7. ISOR-2's main menu

all selected factors worsened (-1), and in field eight the values according to different time intervals are depicted. All data for the twelve months measurements are missing (-9999).

The next step means creating an explanation for the selected patient "D5". From the case-base ISOR-2 retrieves general solutions. The first retrieved one, the PTH factor, denotes that the increased Parathyroid hormone blood level may explain the failure. Further theoretical information (e.g. normal values) about a selected item can be received by pressing the button "show comments". The PTH value of patient "D5" is missing (-9999). From menu field ten the expert user can select further probable solutions. In the example, an increased phosphorus level (P) is suggested. Unfortunately, phosphorus data are missing too. However, the idea of an increased phosphorus level as a possible solution shall not be lost. So, an artificial case can be generated.

The final step means inserting new cases into the case-base. There are two sorts of cases, query cases and artificial cases. Query cases are stored records of real patients from the observed database.

Artificial cases inherit the key attributes from the query cases (field seven in the menu). Other data may be declared as missing, by the update function data can be inserted. In the example, the generalised solution "High P" is inherited, it may be retrieved as a possible solution (field nine of the menu) for future cases.

3.6 Another Problem

Above we described just one of many problems that can arise and that can be solved and analysed by ISOR-2. Another interesting research question is "Does is make sense to begin with the fitness program during the first year of dialyse?" The question arises because the conditions of the patients are considered to be unstable during their first year of dialyse. The question can be expressed in this way "When shall patients begin with the fitness program, earlier or later?" The term "earlier" is defined as "during the first year of dialyse". The term "later" means that they begin with their program after at least one year of dialyse. To answer this question, two groups of active patients are considered, those who began their training within the first year of dialyse and those who began it later (table 2).

Table 2. Changed conditions for active patients

	Earlier	Later
Better	18	10
Worse	6	15

According to Fisher's Exact Test dependence can be observed, with $p < 0.05$. However, it is not as it was initially expected. Since patients are considered as unstable during their first year of dialyse, the assumption was that an earlier beginning might worsen the conditions of the patients. But the test revealed that the conditions of active patients who began with their fitness program within the first year of dialyse improved more than the conditions of those patients starting later.

However, there are six exceptional cases, namely active patients starting early and their conditions worsened. The explanations of them are the additional factors high PTH or a high phosphorus level.

3.7 Exceptional Cases and Outliers

Though statistics is the traditional method for data analysis, other intelligent data analysis methods have become popular, especially for large quantities of data, e.g. for online analytic processing [26]. Such "data mining" methods include artificial neural networks, Baysian networks, decision trees, genetic algorithms, and statistical pattern recognition [27]. However, ISOR-2 is designed for medical studies. Usually, there are just a few cases involved in such studies and the data sets are rather small.

Outlier detection and outlier management are interesting research topics, especially in medicine, because often measuring errors occur due to nurses and doctors being under stress. A strange data value that stands out because it is not like the rest of the data in some sense is commonly called an outlier [28]. However, in the strongest model version in the presented application of dialyse and fitness 22 out of 72 cases (this means about 30 %) contradict the hypothesis (table 1 in section 3.3). Since the threshold of most statistical tests is 5%, in the presented application the contradicting cases should not be treated as outliers.

Furthermore, the idea of ISOR is to support research doctors in their search for reasons why cases are deviating from a research hypothesis. Usually, the doctors are not interested in this question when there are just few outlier cases but when the number of such cases is rather big.

How about the ethical point of view? Patients without health insurance or with serious co-morbidities can become fiscal disasters to those who care for them. Papadimos and Marco [29] presented a philosophical discourse, with emphasis on the writings of Immanuel Kant and G.F.W. Hegel, as to why physicians have the moral imperative to give such "outliers" considerate and thoughtful care. However, the seriously ill dialyse patients should not be blamed if the do not go in for sports actively, because they might feel too week due to their physical condition.

Furthermore, ISOR-2 is a general program applicable on medical studies. In the dialyse and fitness application the patients had the choice to actively participate in the fitness program or to do so rather passively. Usually, in medical studies patients have just the choice between different treatments (often a rather new one and an established one). Beforehand it often cannot be said which choice will lead to an outlier group.

4 Discussion

Since 2003 workshops on Case-Based Reasoning in the health science are regularly organised within the annual CBR conferences. Since then the number of medical CBR systems has already increased and we assume that it will grow further. CBR has been applied in various medical application areas. However, as a tendency we see a slight shift from diagnostic applications to rather therapeutic and other sorts of support systems.

At the last of these workshops [30] a few papers were presented that deal with the consideration of exceptional cases, mainly as deviations from normal situations. Though we are dealing with exceptional cases more explicitly than the other approaches, we believe that this is just the current situation and that a future trend even in medical knowledge-based systems in general might be a more extensive consideration of exceptional cases.

For many years knowledge-based applications have been developed, where medical experts are just considered as knowledge providers. Their knowledge is incorporated in systems that for given input compute results that are presented to the users as output. From our experience, doctors have seldom been happy with such systems, because they especially want to know why and how a system draws conclusions.

We see a new direction emerge, namely conversational systems. The idea is to find solutions within a dialog between the user (a doctor) and a knowledge-based system. Of course, domain knowledge still has to be incorporated in such a system before it starts. Additionally, further knowledge shall be added at run-time. This is a dialog process in which both participants (system and doctor) should stimulate each other. By both ISOR systems presented in this chapter we have shown steps into this direction.

Within the CBR community the first workshop on conversational systems was organised in 2003 [31]. For medical applications, first steps were undertaken even earlier [32]. However, just recently it seems that it is becoming a more popular idea [33, 34].

5 Conclusion

In medical practice and in medical research, doctors very often are confronted with exceptions. Every system that is developed to support doctors shall take this into account. In knowledge-based systems exceptions have to be considered, which may lead to a huge amount of exceptional rules.

In medical studies, statisticians are usually quite happy when a statistical test significantly supports a hypothesis. However, when the number of cases contradicting the hypothesis is rather big, physicians often are not satisfied with significant test results but additionally wish to get explanations for the exceptional cases.

In this chapter, we have presented two systems that deal with both sorts of these situations. The first one, called ISOR-1, especially investigates therapy inefficacy by a dialogue with the user. The second system, ISOR-2, helps to explain cases that contradict a theoretical hypothesis.

First, we have presented ISOR-1, a CBR system designed to help doctors to solve therapeutic problems, particularly to investigate causes of inefficacy of therapies. At present, it deals with two application areas. In the endocrine domain a strong theory exists and therefore usually a lot of solutions are theoretically possible. So, here the task is to reduce a set of possible solutions as much as possible.

ISOR-1 includes different knowledge containers, namely a case-base, a knowledge base, prototypes, and individual bases of patients that reflect their medical histories. Information retrieved from these containers is arranged in form of dialogues. The case-base plays a central role in the dialogue forming process. It serves as a kind of filter when the knowledge base suggests too many possible solutions for the problem (as in the first example). In this situation the most typical cases are retrieved from the case-base. When a solution from the knowledge base is not convincing or when it is hardly adaptable, the case-base may provide better alternatives (as in the example from the psychiatry domain). Generalisations, keywords and references to other knowledge components belong to the case-base. The adaptation program uses them to create dialogues.

ISOR-1 is designed to solve problems from different medical domains. Specific, domain dependant features are mainly attributed to the individual base, because every domain requires a special design of case histories. The knowledge base in ISOR-1 is domain-oriented, but all algorithms and functions are completely domain independent.

Second in this paper, we have proposed to use CBR in ISOR-2 to explain cases that do not fit a statistical model. Since for the dialyse and fitness problem too many parameters exist (nearly 100), the idea is to select three main ones to set up a model and to search subsequently for better parameters to explain the exceptional cases.

The statistical model we presented for ISOR-2 is one of simplest ones but here it is relatively effective because it demonstrates statistically significant dependence between fitness activity and health improvement of dialyse patients. For the strong alternative, the model covers about two thirds of the cases, whereas the remaining one third of them has to be explained by the help of CBR. According to table 1, the stronger the model, the more exceptional cases have to be explained. This means more work for CBR but with a stronger model the confidence in the result of the data analysis increases.

Seven of the exceptional cases seemed to be wrongly classified. After the expert's consultation their classification has been changed. Two kinds of model faults are observed. First, just six months are chosen as relevant time period, whereas for a couple of patients the health condition improved later on. Second, with the choice of qualitative categories, better and worse, it is not possible to assess the amount of the change. So, it cannot be distinguished between small and big changes of the health conditions.

Acknowledgement

We especially thank Prof. Alexander Rumyantsev, Pavlov Medical State University in St. Petersburg, for their close cooperation.

We thank Dr. Monika Mix, Children's Hospital of the University Clinic of Rostock, and Prof. Nikolai Nikolaenko, Sechenov Institute of Evolutionary Physiology and Biochemistry in St.Petersburg, for their data and for their help and time during our consultations.

Furthermore, we thank Professor Aleksey Smirnov, director of the Institute for Nephrology of Saint-Petersburg Medical University and we thank Natalia Korosteleva, researcher at the Institute for Nephrology of St.-Petersburg Medical University for collecting and managing the data.

Resources

A Key Books

Kolodner, J (1993): Case-Based Reasoning. Morgan Kaufmann, San Mateo, CA.

Lenz M, Bartsch-Spörl B, Burkhard H-D, Wess S (eds.) (1998): Case-Based Reasoning Technology. Springer, Berlin Heidelberg.

B Key Survey/Review Articles

Aamodt A, Plaza E (1994): Case-Based Reasoning: Foundational issues, methodogical variations, and system approaches. AI Communications 7 (1): 39-59.

Gierl L, Bull M, Schmidt R (1998): CBR in Medicine. In: Lenz M, Bartsch-Spörl B, Burkhard H-D, Wess S (eds.): Case-Based Reasoning Technology. Springer, Heidelberg Berlin, 273-297.

Schmidt R (2007): Case-based Reasoning in medicine, especially an obituary on Lothar Gierl. In: Yoshida H, Jain A, Ichalkaranje A, Jain LC, Ichalkaranje N (eds.): Advanced Computational Intelligence Paradigms in Healthcare. Studies in Computational Intelligence 48, Springer, Berlin Heidelberg, 63-87

Schmidt R, Montani S, Bellazzi R, Portinale L, Gierl L (2001): Case-Based Reasoning for Medical Knowledge-based Systems. International Journal of Medical Informatics 64 (2-3): 355-367

Nilsson M, Sollenborn N (2004): Advancements and trends in medical case-based reasoning: An overview of systems and system developments. In: Proceedings Seventeenth International Florida Artificial Intelligence Research Society Conference, AAAI Press, Menlo Park, California, 178-183

C Key Journals

Artificial Intelligence in Medicine (Elsevier).

Computational Intelligence (Blackwell)

D Key International Conferences/Workshops

ICCBR – International Conference on Case-Based Reasoning
ECCBR – European Conference on Case-Based Reasoning
AIME – Artificial Intelligence in Medicine Europe
ICDM – International Conference on Data Mining

International Workshops on Case-Based Reasoning in Medicine:
Since 2002 part of ICCBR / ECCBR conferences
International Workshops on Data Mining in Life Sciences:
Since 2006 part of the ICDM conferences

References

1. Marling, C., Whitehouse, P.: Case-Based Reasoning in the care of Alzheimer's disease patients. In: Aha, D.W., Watson, I. (eds.) Case-Based Reasoning Research and Development, pp. 702–715. Springer, Berlin (2001)
2. Bichindaritz, I.: A case-based assistant for clinical psychiatry expertise. Journal of the American Medical Informatics Association, Symposium Supplement, 673–677 (1994)
3. Hirschfeld, R.M., et al.: Partial response and nonresponse to antidepressant therapy: current approaches and treatment options. J. Clin. Psychiatry 63(9), 826–837 (2002)
4. Cuffel, B.J., et al.: Remission, residual symptoms, and nonresponse in the usual treatment of major depression in managed clinical practice. J. Clin. Psychiatry 64(4), 397–402 (2003)
5. Alacorn, R.D., Glover, S., Boyer, W., Balon, R.: Proposing an algorithm for the pharmacological treatment of posttraumatic stress disorder. Ann. Clin. Psychiatry 12(4), 239–246 (2000)
6. Osser, D.N., Patterson, R.D.: Algorithms for the pharmacotherapy of depression, parts one and two. Directions in Psychiatry 18, 303–334 (1998)
7. Working group for paediatric endocrinology of the German society for endocrinology and of the German society for children and youth medicine, 1-15 (1998)
8. Tversky, A.: Features of similarity. Psychological review 84, 327–352 (1977)
9. DeGroot, L.J.: Thyroid Physiology and Hypothyroidsm. In: Besser, G.M., Turner, M. (eds.) Clinical endocrinilogy, Wolfe, London, ch. 15 (1994)
10. Aamodt, A.: Knowledge-Intensive Case-Based Reasoning in CREEK. In: Funk, P., González Calero, P.A. (eds.) ECCBR 2004. LNCS (LNAI), vol. 3155, pp. 5–15. Springer, Heidelberg (2004)
11. Ma, J., Knight, B.A.: Framework for Historical Case-Based Reasoning. In: 5th International Conference on Case-Based Reasoning, pp. 246–260. Springer, Berlin (2003)

12. Gupta, K.M., Aha, D.W., Sandhu, N.: Exploiting Taxonomic and Causal Realations in Conversational Case Retrieval. In: Craw, S., Preece, A.D. (eds.) ECCBR 2002. LNCS (LNAI), vol. 2416, pp. 133–147. Springer, Heidelberg (2002)
13. Davidson, R.J.: Cerebral asymmetry and affective disorders: A developmentalperspective. In: Cicchetti, D., Toth, S.L. (eds.) Internalizing and externalizing expressions of dysfunction. Rochester Symp. on Developmental Psychopathology, Hillsdale, vol. 2, pp. 123–133 (1991)
14. Gelder, M.G., Lopez-Ibor, U., Andeasen, N.C. (eds.): New Oxford Textbook of Psychiatry. Oxford University Press, Oxford (2000)
15. Stein, M.B.: Attending to anxiety disorders in primary care. J. Clin. Psychiatry 64(suppl. 15), 35–39 (2003)
16. Kalinowsky, L., Hippius, H.: Pharmacolological, convulsive and other somatic treatments in psychiatry, Grunee&Stratton, New York, London (1969)
17. Kendall, M.G., Stuart, A.: The advanced theory of statistics, 4th edn. Macmillan publishing, New York (1979)
18. Hai, G.A.: Logic of diagnostic and decision making in clinical medicine. Politheknica publishing, St. Petersburg (2002)
19. Bichindaritz, I., Kansu, E., Sullivan, K.M.: Casebased Reasoning in Care-Partner. In: Smyth, B., Cunningham, P. (eds.) EWCBR 1998. LNCS (LNAI), vol. 1488, pp. 334–345. Springer, Heidelberg (1998)
20. Prentzas, J., Hatzilgeroudis, I.: Integrating Hybrid Rule-Based with Case-Based Reasoning. In: Craw, S., Preeece, A. (eds.) ECCBR 2002. LNCS (LNAI), vol. 2416, pp. 336–349. Springer, Heidelberg (2002)
21. Shuguang, L., Qing, J., George, C.: Combining case-based and model-based reasoning: a formal specification. In: Seventh Asia-Pacific Software Engineering Conference, APSEC 2000, p. 416 (2000)
22. Corchado, J.M., Corchado, E.S., Aiken, J., et al.: Maximum likelihood Hebbian learning based retrieval method for CBR systems. In: Ashley, K.D., Bridge, D.G. (eds.) ICCBR 2003. LNCS, vol. 2689, pp. 107–121. Springer, Heidelberg (2003)
23. Rezvani, S., Prasad, G.: A hybrid system with multivariate data validation and Case-based Reasoning for an efficient and realistic product formulation. In: Ashley, K.D., Bridge, D.G. (eds.) Proceedings International Conference on Case-based Reasoning, ICCBR 2003, pp. 465–478. Springer, Berlin (2003)
24. Arshadi, N., Jurisica, I.: Data Mining for Case-based Reasoning in high-dimensional biological domains. IEEE Transactions on Knowledge and Data Engineering 17(8), 1127–1137 (2005)
25. Davidson, A.M., Cameron, J.S., Grünfeld, J.-P., et al. (eds.): Oxford Textbook of Nephrology, vol. 3. Oxford University Press, Oxford (2005)
26. Codd, E.F.: Providing OLAP (On-Line Analytic Processing) to User-Analysts: Am IT Mandate. E.F. Codd and Associates (1994)
27. Piatetsky-Shapiro, G., Ffrawley, W.J.: Knowledge Discovery in Databases. AAAI Press/ The MIT Press (1991)
28. Cheng, J.G.: Outlier Management in Intelligent Data Analysis. PhD Thesis (2000), `http://www.dcs.bbk.ac.uk/research/recentphds/gcheng.pdf` (last visited May 2008)
29. Papadimos, T.J., Marco, A.P.: The obligation of phyicians to medical outliers: a Kantian and Hegelian synthesis. BMC Medical Ethics 5(3) (2004)

30. Bichindaritz, I., Montani, S.: Workshop – Case-Based Reasoning in the Health Science. In: Wilson, D.C., Khemani, D. (eds.) ICCBR- Workshop Proceedings, Belfast, pp. 257–368 (2007)
31. Aha, D.: Workshop – Mixed-Initiative Case-Based Reasoning. In: McGinty, L. (ed.) ICCBR 2003 Workshop Proceedings, Trondheim, pp. 127–209 (2003)
32. McSherry, D.: Interactive Case-Based Reasoning in Sequential Diagnosis. Applied Intelligence 14, 65–76 (2001)
33. McSherry, D.: Hypothetico-Deductive Case-Based Reasoning. In: Bichindaritz, Montani (eds.), pp. 315–324 (2007)
34. Vorobieva, O., Rumyantsev, A., Schmidt, R.: Incremental Development of an Explanation Model for Exceptional Dialysis Patients. In: Bichindaritz, I., Montani, S. (eds.) Workshop on CBR in the Health Sciences, ECCBR 2006 Workshop Proceedings, pp. 170–178. University of Trier Press (2006)

Chapter 9

Security in Medical Telediagnosis

J.-B. Aupet, E. Garcia, H. Guyennet, J.-C. Lapayre, and D. Martins

LIFC - EA 4269, Université de Franche-Comté,
16, Rte de Gray - 25030 BESANCON Cedex, France

Abstract. Telemedicine is the process by which electronic, visual and audio communications are used to support practitioners at remote sites with diagnosis and consultation procedures, such as remote clinical examinations and medical image transfers. Telemedicine is legally regulated by laws and constraints regarding the access of data contained in Personal Medical Files. These requirements are given by international entities such as IHE. They often specify the required functionalities to meet requirements of information exchange in the medical field. These functionalities are presented in this paper and outline some technical examples for their implementation in each level: authentication, communication, data storage, tracking and patient identification.

Keywords: Authentication, Biometric, Cryptography, Medical Software, Medical Data, Remote Telesurgery, Secured data, Secured Network, Telediagnosis, Telemedicine, Tracking.

1 Introduction

Over the past 10 years, the concept of working remotely (teleworking) has been in rapid development. This phenomenon is due in large part to the parallel growth in high performance networks and processors, but also due to the improvement of security in these systems. Teleworking is used in various ways such as distant learning, remote maintenance and even telemedicine. The CARTOON (Collaborative Architecture Distributed Algorithm Optimization and Network) Group at the LIFC (Franche-Comté Computer Science Lab. in France) has acquired a good experience on collaboration management for medical e-diagnosis. The TeNeCi project(Collaborative Teleneurology, 2006) and the Decopreme project (Precocious Collaborative Screening of Melanomas, 2007) are parts of INTERREGIII program in collaboration with Swiss partners (Vaud University Hospital at Lausanne, and EPFL Lausanne).

Telemedicine is generally used in a non-acute setting for patient monitoring or education and has only recently been introduced into emergency care. Telemedicine can be defined as the use of telecommunication technologies to provide medical information and services. It is the process by which electronic, visual and audio communications are used to support practitioners at remote

G.A. Tsihrintzis and L.C. Jain (Eds.): Multimedia Services in Intel. Environ., SIST 3, pp. 201–226.
springerlink.com

sites with diagnosis and consultation procedures, such as remote clinical examinations and medical image transfers.

For this kind of application, graphic interface and additional tools must facilitate actors capacity to disregard distance and time in order to reconstitute a virtual examination room. Software and network architecture have to be optimal to improve interactivity and fault tolerance. Our aim is to obtain a secure environment to exchange medical data, diagnoses and opinions.

This chapter is composed of four sections after this introduction. In section two, we present the security requirements for telemedicine and telediagnosis. Section three exposes methods of authentication such as cryptography with certificates or SSO modalities and other solutions. Section four presents a secured medical data storage then in the next section a state of art in secured data transfer. Finally the last section describes a new secured telemedicine software which shows solutions to implement security in this kind of application.

Definition of Telemedicine and Telediagnosis
It is interesting at the beginning of this chapter to define the context of our study. ((SET 2009), (ISfTeH 2009)).

The rapid progress in telecommunications in recent years has allowed the practice of telemedicine to grow. Telemedicine means "*practice of medicine remotely by means of telecommunication*". It is multifaceted:

Telesurveillance
For example, the tools to supervise patients at home using internet or GPRS. The EPI-MEDICS project and the UR-SAFE project try to propose solutions for remote monitoring of patients. The idea is to give to the patient a mobile device that can retrieve and transmit data such as blood pressure or heartbeat (Mailhes 2003), (Gouaux 2002).

Telesurgery
In September 2001 the first telesurgery was realized by Professor Jacques Marescaux (Marescaux 2001), (Marescaux 2002) specialist in digestive surgery. He operated from New York a patient who was at the Strasbourg hospital. Another form of telesurgery is teleassistance for operating gesture and remote handling. In this domain, another projects make it possible to share patient's medical records in real time for in-depth analysis and enhanced preparation of optimum surgical strategies : for example Argonaute3D (Le Mer 2004).

Teleconsultation
Teleconsultation is used in different fields of medicine (such as neurology (Garcia 2005) , psychiatry (Worth 2003), or dermatology for precocious collaborative screening of melanomas (Elmarzouqi 2007), ...). It is a possibility to allow a practitioner a distant help (Mlabs 2009) with a Multimedia Conferencing System for example), by a senior practitioner for example. This collaboration between practitioners is not new, it has existed for a long time, but the means were different: in general, information exchange was done by phone. For example, on Figure 1 a patient arrives to the hospital and is admitted to the emergency service (a): different data are collected (electroencephalography, cerebral scan...).

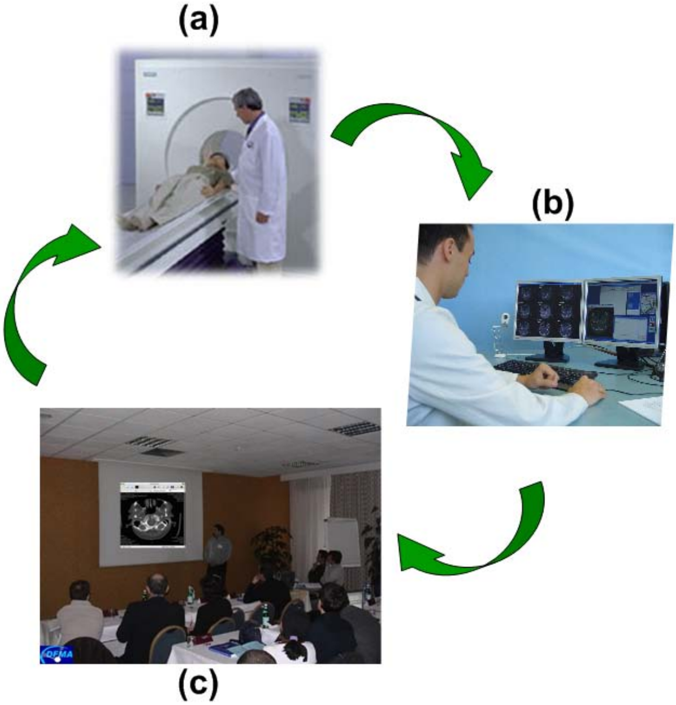


Fig. 1. Telediagnosis example: (a)Emergency, (b)Remote Practitioner, (c) Telestaff

Now it is possible, in real time, to have a distant help with a practitioner (b) or with a staff of practitioners (c).

Teleexpertise
The idea is now to allow remote expertise. Many projects are developed for the tele-use of echography (figure 2): the OTELO project (*mObile Tele-Echography using an ultra Light rObot) (OTELO 2004), the ESTELE project (*Expert System for Tele Echography), the ARTIS project 2 or the TERESA project in (Vieyres 2003). These projects propose a solution to make an echography remotely: for example TERESA is a tele-echography project that proposes a solution to bring astronauts and remotely located patients on ground quality ultrasound examinations despite the lack of a specialist at the location of the medical act.

Remote Vocational Training
During a practitioner's career, it is necessary that the knowledge is put up to date regularly: medicine evolves. Vocational training is not necessarily given in hospitals and it is very interesting to minimize the business trips of senior practitioners: the remote teaching is one of the solutions.

E–health Network
These networks are developed to encourage access to the medical attentions, coordination, continuity or interdisciplinary of care, notably of those that are specific to some populations, pathologies or sanitary activities. They assure an adapted care to the needs of the patient: from the education to health and prevention to the diagnosis. They can contribute to the actions of public health.

In all of these telemedicine domains the security is an important point. Access is remote by the networks and computer systems. And therefore problems of security appear: connection to the system, securization and transfer of data.

Fig. 2. ARTIS project: Remote Tele Echography

2 Security Requirements for Telemedicine and Telediagnosis

Many technologies are now available to secure electronic information in all domains. Telemedicine has its own rules depending on the country in which they applied.

We can find different directives in the laws of the different countries; we can also find requirements given by international entities such as IHE. They often specify the functionalities to meet requirements of information exchange in the medical field. Some countries also require a certification of the telemedicine applications to allow their deployment.

These functionalities are presented in this paper and outline some technical examples for their implementation. Among these features, we can find:

The authentication

The first requirement concerns the methods of authentication. The aim is to ensure that a user is allowed to enter the system and manipulate confidential data. It also requires the ability to verify if a specific user (e.g. a physician) can access the data of a particular patient according to the rules applied in the country. A patient for example can choose his doctor and ask to be allowed to have access to only part of the file.

In addition, the node (computer) used by the user also needs to be authenticated in different applications, for example, for applications that need to connect to high security systems, such as PACS (Picture Archiving and Communication Systems). In this case, only identified stations are allowed to send and retrieve data from the system. Even if the user is authenticated and authorized, he can't perform transfer with these systems without being on allowed nodes. Indeed, a great number of medical systems are in shared access, and the identity of the computer is more important than the identity of the user. The physical access to these systems is often controlled.

Authentication of users and nodes can be managed by many technologies, including login/password, card, RFID card, biometrics... It is frequently recommended to use certificates in most cases based on X509 to encrypt communication.

As the authentication becomes more and more required, a health information system composed of several applications makes it heavy for the user to launch. Indeed, he will have to authenticate several times, with several means, therefore, it is recommended to use a unique mean of authentication (based on a RFID, on a login/password with a LDAP repository...) with a policy avoiding multiple authentications. It can be a RFID which automatically authenticates the user when he launches an application, or an SSO system (Single Sign On) asking the login/password once and using it for each application.

The data security

All data has to be secured when stored and during communications. Two means can be used: the use of a VPN and data encryption. The second one may be used to avoid the installation of a VPN technology. Encryption provides security for information while in transit. Different technologies can be used for encryption: DES, 3DES, AES... A common way to transfer data is to use SSL or TSL connections. Medical systems can also have to meet integrity control requirement. In this case the use of digital signature is recommended. This allows the system to check the identity of the sender and ensure that data has not been modified by another user. Anonymization must also be respected. There exist cases where the identity of a patient must be automatically deleted by the system like the use of medical data for scientific or teaching use. Generally, images belong to the hospital, and if a user wants to send them to someone out of the hospital, telemedicine system has to support this functionality.

The requirements have to be met without making it too heavy to deploy and use telemedicine applications (they have to be compliant with common security functionalities such as proxies and firewall).

The tracking
Medical applications need to offer audit trail tools that track all exchanges of medical data. IHE ATNA specifies a standard way to log these events. Reliable Syslog (RFC 3195) is the recommended transport protocol for audit messages. Events and audit trail messages are based on IETF, DICOM, HL7 and ASTM standards. For example, one must log successful connections, as well as successful and failed transfers... Such a system must allow the administrator to verify who has used the system and who has accessed a patient file...

Another issue here is the recording of actions during a work session. Let us take the example of a collaborative diagnosis. All or part of the actions could be recorded. In such cases, the telemedicine application has to provide the functionalities to keep a guaranteed information such as: who has performed this action... These records could be used for teaching, knowledge management, as well as for responsibility research.

Other specific requirements
Patient Identification: Many countries do not have a unique patient ID for different sites, thus another important requirement for a telemedicine system is to ensure that a patient is highly identified (on one site or on different sites in case of transfer) in the system to avoid any error. PIX standard is often used for such a feature.

Storage: Storage is also a large issue in data security. Two aspects must be taken into account:

– to protect data from intrusion: this is made by encryption of data and by putting the data on protected servers
– to protect data from loss: this is ensured by using replicated servers with backup functions and using a specific life cycle of data (recent data is stored on server hard disks, and then stored on other type of supports according to their age).

Bandwidth: Some countries have a separated network for the health system. To deploy an application on such a network, a certification could be required. Software is tested for security issues but also for bandwidth consumption and avoids deploying an application that could break down the network. It is a feature to take into account to obtain this type of certification.

3 Methods of Authentication

Telemedicine is legally regulated by laws and constraints regarding the access of data contained in Personal Medical Files. Telemedicine systems must formerly respect these obligations. Patients could have access to their personal files, know what information it contains, who has entered the information and have the ability to rectify or delete saved information. To ensure these requirements, the system must implement strong and sure authentication methods and traceability mechanisms.

Authentication is the process of determining whether someone or something is, using a method approved by the asked information provider (application, database, internet server…). Authorization rights were also managed by the authentication scheme. With the increase of data transmission on computer networks, the rate of sensible data transmitted is becoming important. So the authentication is required ensure the identity of each actor on the system. This concept has been studied since the early days of computing (Lamport 1981), networking (Needham 1978) and distributed system (Lampson 1992).

In Telemedicine softwares, patient's files should be accessible only for allowed practitioners and only with patient agreement. When the user (or machine) has been authenticated, secured data can be exchanged and the access to the allowed services according to the granted rights will be authorized. There exists different methods of authentication but each has advantages and inconveniences.

The authentication process is a compromise between the measure of risk and the computing time to authenticate, but also depend on the computing possibilities of the terminal, for example, mobile networks (MacDonald 2008), PDA, 3G Smartphone. For high risk systems such as applications and information exchanges, an authentication that accurately confirms the user's digital identity is chosen. For a low risk application, where the confirmation of the digital identity is not as important from a risk perspective, a method that is more prompt is chosen. Authentication methods differ by using something the user knows such as password (Lamport 1981), something the user has such as security token (Schwiderski-Grosche 2006) or directly a part of the user itself, biometrical properties (Tuyls 2004), (Bernecker 2006), (Ruud 2001). These methods will be enumerated and we will show how to use and combine them to make sure that no one can usurp a protected identity.

3.1 Authentication Possibilities

Password Authentication

This is the commonly used authentication method, but is also the least secure. Password authentication verifies the identity of the user with user id and password, in order to login. The security level is set by password management including configuration of password length, type of characters used and password duration. The ability to easily crack passwords has resulted in high levels of identity theft. The risk of using passwords has resulted in deploying a layered security strategy. Password authentication method is used only for low risk applications and other forms of authentication are required for higher risk applications.

Lightweight Directory Access Protocol (LDAP) Authentication

Most enterprises use Lightweight Directory Access Protocol (LDAP) directories to handle the centralized authentication. LDAP directories, such as Active

Directory, Sun One Directory, Novel e-Directory and other vendors, provide a low cost way of doing fast identity look-ups and authentication as compared to traditional databases. Today it is common to use virtual LDAP directories to quickly integrate the identity and authentication information contained in one or more databases and/or other LDAP directories. The use of these directories is a critical piece of the identity infrastructure that leads to integrating access control.

PKI Authentication

Public key infrastructure (PKI) authentication, is another way of doing identity authentication. An identity is given a digital certificate by a Certificate Authority (CA). Then presented during the authentication process to verify an identity versus his own key. The level of authentication trust varies for digital certificates depending on the level of identity verification done during the identity registration process as well as the digital certificate revocation process. Digital certificates are becoming more important to authentication and verification of an identity in single sign on systems, document management systems and in web services.

Security Token Authentication

Security token authentication, such as RSA secureID tokens, are used to make authentication of an identity. During the login process, or if required by a single sign on system for a higher risk application, the user is required to enter in the numbers appearing on the token screen along with their id (Figure 3). Since the numbers change randomly to the user viewing the screen (but is understood by the central authentication server), there is a higher degree of trust associated with this form of authentication. Unfortunately, operating costs for security authentication tokens are higher than the use of password and id since they have to be physically issued, replaced and recovered.

Smart Card Authentication

Smart cards are another form of authentication token. Often they contain a digital certificate as well as additional identity attribute information. Information can be stored in an RFID or Mifare no-contact chip or in a microchip or magnetic stripe systems (Figure 4). Smart card authentication is becoming wide spread. The same smart cards used in an authentication process are now commonly used for access control mechanisms to enter physical facilities, buildings, floors and rooms. This type of card is commonly used in medical infrastructures to secure rooms and authenticate users, therefore it can be used to authenticate users on the information system.

Network Authentication

Network authentication is the process of granting a user (or node) the ability to authenticate a network as well as their authorization. Almost all network authentication systems are now LDAP based, including Microsoft 2000, Linux, Solaris, AIX and HPUX. Many mainframe authentication systems such as RACF

Fig. 3. Token Generator

Fig. 4. Smart Card

are now LDAP enabled. Network authentication is commonly assured by password, but most enterprises replace or add to this system another authentication modality.

Biometric Authentication
Biometric authentication is the process of taking a "piece of you", digitizing it to formerly authenticate the user on the system, then he can access his personal directory or database. Typical types of biometric authentications include digital finger prints, hand scans, retina scans, digital signature scans or vein scans (figure 5). All these technics are available and can be used for authentication. They can compared with several criteria such as False Accept Rate and False Reject Rate (table 1) and the Equal Error Rate that is a result of both as shown (figure 6). Biometrics are commonly used as part of an array of authentication methods used in information systems.

Fig. 5. Vein Scanner

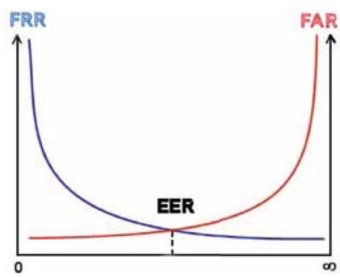


Fig. 6. Equal Error Rate obtention

3.2 SSO Modalities

Single Sign On (SSO) is the ability of a user to enter the same id and password to logon to multiple applications within an information system. As passwords are the least secure authentication mechanism, single sign on has now become known as reduced sign on (RSO) since more than one type of authentication mechanism

Table 1. Comparative table of biometrical methods

Part	Finger	Voice	Iris	Face	Vein
Type	Physical	Behavioral	Physical	Physical	Physical
Method	Active	Active	Active	Passive	Active
Equal Error Rate	2 - 3,3%	$< 1\%$	4,1 - 4,6%	4,1%	1%
Failure to Enroll	4%	2%	7%	1%	0.02%
Nominal False Accept Rate	2.5%	$< 1\%$	6%	4%	0.0001%
Nominal False Reject Rate	0.1%	$< 1\%$	0.001%	10%	0.1%
Liveness Aware	Possible	Yes	Possible	Possible	Yes
System Cost	Average	Low	Very High	High	High

is used according to risk models. Single sign on can also take place between enterprises using federated authentication. For example, a business partner's employee may successfully log on to their system. When they click on a link to your application, the business partner's single sign on system will provide a security assertion token to your information system using a protocol like SAML, Liberty Alliance, WS Federation or Shibboleth. Your company's SSO software receives the token, checks it, and then allows the business partner's employee to access your business application without having to sign on.

For example, in a company using SSO server, users log on with their id and password, gaining them access to low risk information and multiple applications such as the business portal. However, when the user tries to access higher risk applications and information, like a payroll system, the single sign on server requires them to use a stronger form of authentication. This may include digital certificates, security tokens, smart cards, biometrics or combinations thereof.

Single Sign On benefits are:

- Ability to enforce uniform authentication and/or authorization policies across the information system.
- End to end user audit sessions to improve security reporting and auditing.
- Removes application developers from having to understand and implement identity security in their applications.
- Usually results in significant password help desk cost savings.

Since the internet is stateless, this means that the single sign on software must check every request by the user's browser to see if there is an authentication policy pertaining to the resource or application the user is trying to access. In a medium to large company or medical center, this means that every time the user clicks on a different URL, there is traffic between the user's browser, the web or application servers and the security server. This traffic can become large and cumbersome from a performance perspective. Therefore, most modern single sign on systems use LDAP (Lightweight Directory Access Protocol) directories to store the authentication and authorization policies. The LDAP directories are made for high performance lookups thus addressing the high traffic load. Further, the LDAP directories are often the source for the single sign on system to authenticate against. Single sign on systems in medium to large information systems can become a single point of failure if not properly designed. If the single sign on system goes down but the applications remain up, no user can access any resource or application protected by the SSO system. Many companies have experienced this painful condition resulting in productivity loss. Therefore, it is essential that your single sign on system have a good and well tested failover and disaster recovery design.

Finally, single sign on systems requires good identity data governance. Security features being offered by the single sign on system is only as good as the underlying identity data. Thus it is critical that all identity data have good, quick business processes that pick up on any change to the identity such as new identity creation, identity termination or role changes. Without this, SSO systems are vulnerable to creating security holes.

4 Secured Data Storage

As the population of the world is growing quickly, and as new medical diagnostic technologies are developed, the traditional archival system produces several major problems in accommodating the increased demand. Consequently, more and more physical space will become necessary to store medical files, and it will become increasingly less efficient and more difficult to access and retrieve a particular medical file (Zhang 2005). In recent years, there has been an evolution of technology access and storage, and especially with increased legal obligations.

Two categories of staff have access to patient data: administrative staff, and nursing staff. Their ways to access archived data are radically different. The nursing staff is a population which is very difficult to secure manner access: open access to workstation, shared access by several people, with the imperatives of access speed incompatible with strong authentication techniques. In addition this medical data storing can be entrusted to companies outside the hospital that will also secure. All this information is new wealth for institutional care but it must be always available and reliable.

With the rapid development of computer and network technologies, it is possible to share and access to data on grid storage devices. In (Zhang 2005), a web-based medical information archive system enables storage, maintenance, sharing, updating, and retrieval of medical data based on the existing Internet facilities.

4.1 DICOM Standard

Digital Imaging and Communications in Medicine (DICOM) (Pianykh 2008) is a standard conceived over 20 years ago for storing and transmitting medical data. This standard enables different DICOM modality such as scanners, MRI, PET, to communicate over a TCP/IP network and then exchange and store medical data into a picture archiving and communication system (PACS). The DICOM C-STORE service is used to send data to a file server (e.g. PACS). However, there is no guarantee that the data has been archived. For this reason, the Storage Commitment was introduced in addition to storage services to ensure that data received on archiving has been taken into account. It explicitly takes the responsibility of good data archiving. For example, a single workstation can implement a reception service tomography images in order to display them and then delete them after use. Another example is a scanner which produces series of medical images. Once produced, these images are sent to a server for archiving. After sending data, the scanner asks for a storage commitment. If the commitment succeeds, the scanner can delete data from its memory. Nowadays, more and more DICOM devices offer this service confirmation archiving, although this storage commitment is purely an administrative problem. Indeed, anyone can log in (username and password) on a storage device and delete some medical data.

DICOM applications offer strategies to try to tackle this problem such as lockers or no-delete flag but none of them really solve the problem.

4.2 Security Requirements for Medical Data Storage

All medical data are considered as sensitive to preserve patient privacy. To read the content of an image, a user needs to be authorized both to access the file and the encryption key. Data anonymization prevents the exposure of most sensitive data to unauthorized users. Several services are required to secure the storage of medical data (Montagnat 2008):

File access control
Control access is most often used to assure a higher degree of security of server. This technique allows connecting any given port of the server and accept particular services and reject suspicious request for access. In the future, medical image access will use digital videos, smart phones and increasingly lighter mobile systems, consumers expect the same level of mobility from digital medicine.

File anonymization
The patient data are confidential and should not be open to the general public. DICOM anonymization software keeps a list of 18 major confidential attribute types (Name, Social security numbers, account numbers) and removes them from DICOM files.

File encryption
Encryption is the process of changing the format of the data to protect its content. Storing the data in encrypted form considerably reduces the risk of disclosure. The encrypting medical files before emailing them is a great way to ensure that identity thieves and others do not have access to those sensitive records. The DICOM standard permits to encrypt the data to verify the data origin and the data integrity.

Secure access to data
Data should be locally stored but also faraway on a remote server. For example in teleriadiology, the images made in one location can be interpreted and read in a completely different one. The only way to transmit images faster implies better networks and the use of lossy image compression. In the case of application email, direct connection to an email server from a remote DICOM application may be rejected for security reasons or overflow mailbox quotas.

4.3 Distributed Storage of Medical Data

Distributed architecture allows users from different organizations to share storage capacity and data. However distributed systems improve the number of users, the number of networks and the number of the storage devices. Because of the cross institution nature of distributed application communications, distributed storage has specific security needs. Network security is a hard-to-define paradigm in that its definition varies with the different organizations which implement it. Security is defined by the policies that implement the services offered to protect the data. Since all medical data are stored on electronic devices, it is important to establish a host-based authentication approach. In (Montagnat 2008) authors propose software architecture to implement a medical data management. This system provide access to medical sources for Grid services and users while taking into account the constraints related to clinical practice. In (Seitz 2005) authors present grids as architecture for medical image processing and health-care networks. The goal is to create a scheme that allows the storage and access to encrypted data on grid storage devices.

5 Secured Data Transfer

Telemedecine's applications need a secure data transfer in a network. This transfer has to ensure four requirements.

Data Confidentiality
The network has to ensure that transmitted data is confidential and respects the confidentiality's rule for a patient. An external person of the network should not be able to read the information exchanged. To secure data, the exchange in the network is to be encrypted by using a private key secret that only intended receiver's posses.

Data Integrity
Data in the network should not be altered by the communication. We have to ensure that nobody can enter in the network and modify data. Consequences could be the death of the patient. This is why the data integrity is probably the most important thing in the security of a telemedecine's application.

Data Freshness
In medical area, the freshness of data is very important. By freshness, we mean that data is recent or not. If we do not pay attention on the freshness of data, an attacker could replay old data. For example, someone can resend critical information about a patient (such as an old decline in insulin) to mislead the nurses, therefore they give insulin to a patient, who does not need, leading to disastrous consequences.

Availability
Data should always be available. It is vital that information can be transmitted without interruptions. We have to ensure that data cannot be stopped. In telemedecine's applications, every second is important and data has to be receiving in time.

In the following parts, vulnerabilities of data transfer in telemedecine's networks will be shown, such as wired networks, wireless networks and more specifically in mobile ad-hoc networks (MANET) and wireless sensor networks. Then solutions for the attacks in the different networks are presented, which can be used to secure data transfer in telemedecine's networks.

5.1 Vulnerabilities of Telemedecine's Networks

Attacks in wired networks

Wired networks are obviously the most secure networks. However, Internet is often used in telemedecine to transfer data from one hospital to another. It is known that the Internet is not the most secure network. It is very important in telemedecine to have a security policy against these attacks and to know these vulnerabilities. Here is a short list of current attacks in wired networks.

Packet sniffing
If data is not encrypted, packet sending on the network or in Internet can be sniffed. Then an attacker can rebuild the entire data and read data information. This attack is a passive attack, but the data confidentiality is broken, and the attacker could know private information of patients.

Denial of service attack
This is one of the most known active attacks. The aim is to saturate by different ways a computer or a server. In telemedecine this kind of attack endangers the availability of the network. For example, a doctor who uses a telemedecine's application to operate a patient with a real time server. The server sends and receives video information. If an attacker makes a denial of service attack against this server, the doctor could not finish the operation.

Man in the middle
The attack of the man in the middle consists of an active eavesdropping in which the attacker makes independent connections with the victims (such as two

persons communicating in two different hospitals) and relays messages between them. They believe that they are talking directly to each other over a private connection when in fact the entire conversation is controlled by the attacker. With this attack, an attacker can listen, modify or delete data.

Replay attack
An attacker uses a sniffer to catch packets sent by a computer. Even he cannot read data, he can masquerade the network by replaying this data in the network later. The data will be compromised.

Attacks in wireless networks

Telemedecine's applications use wireless networks more and more. Wireless offers more mobility and possibilities. A wireless telemedecine's application can be used for a field hospital in a war or catastrophic area. However a wireless network is less secure than a wired network, because the medium is wave. This medium allows the possibility for other specific attacks. Obviously in wireless networks, the same attacks are found as in wired network. Here are next the two specific attacks of wireless network.

Eavesdropping
This passive attack consists of listening to the network to intercept information on the network. This attack is easy, if data is not encrypted and the medium is radio wave. Since this attack does not modify the data, it is difficult to detect it. It attacks directly the confidentiality of data sent by the network. An attacker could read data of patients. Even if the data is encrypted, the network can be eavesdropped, if the network uses a crypted protocol easy to break. For example the wired equivalent privacy (WEP) protocol, which is used in some hospitals, can be easily uncrypted in less than 10 minutes with a basic laptop.

Radio jamming
An attacker sends radio waves at the same frequency that it used by wireless networks. The network cannot communicate if the transport medium is flooded by radio interferences and availability of the network is broken.

Attacks in MANETS

Mobile ad-hoc network (MANETs) is a specific wireless sensor network. An example of MANET in telemedecine is a network with a PDA for each doctor, who communicate together and send information to one or more computers. In MANET, each computer or mobile host (PDA, sensor, etc...), called nodes, communicate with their neighboring nodes and not directly with the base or sink. In a wireless network, hosts communicate directly with a base station, which sends data to the Internet or sends information to another computer. But in MANET, a node, which will communicate information to the base station or a specific node, has to find the lower path to the receiver and send data to its neighbors. The neighbors will then send data to their neighbors and finally data will be received by the base station or the receiver. The particularity of this kind of network creates the possibilities of other specific attacks.

Black Hole Attack

The black hole attack consists at first the insertion of a malicious node in the network (Karlof 2003) (laptop, sensor, pda, etc..). This malicious node, will change routing tables, in several ways, to force a maximum of neighboring nodes to send data to it. After that, like a real black hole in space, all recovered data will never be sent back by the malicious node.

Selective Forwarding (Grey hole attack)

This is a variant of the black hole attack (Karlof 2003). Like in the black hole attack, an attacker will insert a malicious node in the network and this node will change the routing to capture data around it. Unlike the black hole, the attack of selective forwarding relays information. For example, the malicious node will relay all information concerning the routing and it will not relay data, which is critical. That is why, this kind of attack is more difficult to detect than the black hole attack. If the malicious node works normally, it cannot easily be detected.

Slowdown

An attacker can use some malicious nodes to slowdown the network. It can use a selective forwarding attack to do it. This slowdown may be crucial if the network sends critical information of a patient. The information will be slowed and the patient could die.

Wormhole

This attack needs to insert in the network at least two malicious nodes (Hu 2006). These nodes are connected by a powerful connection such as a wired liaison or a powerful wireless signal. This attack wrongs the other nodes of the network by the distance between the two bad nodes, and proposes a quicker route. Generally the routing protocols search the route with the shortest number of hops. In a wormhole attack, the two malicious nodes propose to achieve a distant position with an unique hop. This possibility will wrong other nodes on the real distances that separate the two malicious nodes. The nodes will choose this shortest route for sending their data, and thus send their information to the malicious nodes. The attacker can then catch some information of the network.

Sinkhole

A malicious node will attack the data directly, which circulate near the sink or base station, as the sink is the point at which there is the maximum of data on the entire network (Karlof 2003). To do this attack, the malicious node will offer the quickest route to reach the sink, using a powerful connection. Nodes, near the malicious node, will send data for the sink to it. All information, which is sent from these nodes to the sink, may be captured by the attacker.

Sybil attack
A Sybil attack (Newsome 2004) is a malicious sensor which is masquerading as multiples sensors. It will modify the routing table, which will be wrong.

Attacks in WSNs

Wireless sensor networks or WSNs are specific MANETs. They can be used by example in telemedecine with one or more patient wearing sensors to know their health status such as the CODE BLUE project (Malan 2004). The specificities of WSNs are a low power of calculation and a limited energy. These two differences are used by attackers to endanger the network.

Sleep deprivation torture
An attacker sends many messages or asks calculations to a sensor. The aim is to prevent the sensor to sleep to consume his energy until the sensor becomes out of order. This active attack prevents a sensor to sleep in different ways (Stajano 1999). If the sensor cannot sleep, it will consume its battery and be out of order. In telemedecine, if a sensor is out of order, patient data is no longer read, and the network will not receive vital information.

Flooding
An attacker will use one or many malicious nodes or something else with a powerful signal, sending some messages regularly some messages into the network, flooding it. This is an active attack such as denial of service and the consumption of the energy of nodes in the network.

Infinite loops
An attacker can use two or more malicious nodes to send infinitely packets on the network. As these messages will be endlessly sent by the network like a ping-pong game, sensors will consume their energy and the network will saturate.

Pace maker
Pace maker is a specific sensor which uses electrical impulses to steady the beating of the heart. The pacemaker can be controlled by an external device to change the frequency of beats, however the connection is not secured. In (Halperin 2008), they show that it is possible to make an attack to derange the pacemaker. This attack shows us the problem of data transfer in a sensor without security and consequences.

5.2 To Secure Data Transfer in Telemedecine Network

Secure wired networks

Solutions of security for a telemedecine application in wired networks are not different from solutions existing in other applications that need security. The most important thing is to encrypt data. We have to use existing protocols to do it, such as Secure Sockets Layers (SSL) or Secure Shell (SSH). Both use public key infrastructure and ensure considerably data transfer, principally the data confidentiality and data integrity. The aim is to create a virtual private network (VPN) to secure data transfer. To counter denial of service attacks in telemedecine's application, which use a server for communication, it can be better to have two or more servers. If a server is attacked by denial of service attacks, the other servers can take over and ensure the availability of the network. For data freshness in data transfer, a telemedicine's application has to insert time information in encrypted packets. This time information should be encrypted too, to counter the replay attack. However, some other solutions are given by possibilities of trust computing and the trust computing group, by using a trusted platform module in the computer of the network (Abendroth 2006) to increase the security of the network.

Secure wireless networks

Wireless networks can use the same solutions as wired network to secure telemedecine's application, but they need also to secure data transfer with radio waves. The principal point is to not use WEP. WEP uses the RC4 algorithm with a secret key of 40 or 128 bits. This key is too short, and it is too easy to find it for an attacker. WIFI Protected Access or WPA must be used, as it uses the TKIP algorithm or Temporary Key Integrity Protocol. The TKIP randomly generates a key and changes it many times by second. Moreover WPA uses a authentication server RADIUS for identify users and define access rights in the network. With this solution, it is harder for an attacker to eavesdrop or to enter in a wireless telemedecine's network.

Secure MANETS

In MANETs, WPA2 is principally used, which is a version of WPA supporting MANET, to secure data transfer with radio wave. To counter wormhole attacks some solutions are proposed such as a packet leash protocol (Hu 2003), the SECTOR mechanism (Čapkun 2003) or using directional antennas. The fact to use directional antennas limits the available listen area, for an attacker. Moreover (Seung 2001) presents a security-aware ad hoc routing protocol (SAR), a solution against black hole attacks. Other specific solutions are discussed in (Hu 2004).

Secure WSNs

The specificities of WSN are a low power of calculation and of limited energy, which require to find specific security's solution for WSNs (Martins 2008). For example in WSN, we cannot use public key infrastructure, because a sensor cannot do calculations of this security's solution (Piotrowski 2006). We can only use symmetric secret key to encrypt information. This encryption is easy to break. This is actually why security in WSNs is a big way of research. Other secure solutions need to be found, however there are two popular secure protocols, SPINS (Perrig 2002) and TinySEC (Karlof 2004), which can be used to secure a telemedecine application.

SPINS is a protocol based on two blocks of security SNEP and μTESLA indent SNEP uses two security mechanisms. The first is to encrypt to ensure the confidentiality of data and the second is to use a code authenticity of messages MAC (Message Authentication Code) to ensure authentication and data integrity between two entities.

μTESLA uses a symmetrical authentication linked to an asymmetrical method where the symmetrical keys are disclosed over time.

TinySEC is a security package integrated in the operating system TinyOS.(TinyOS 2009).

The aim of this link layer is to detect unauthorized packets while they are injected for the first time into the network, to prevent their spread in the network that would generated by communications, a loss of energy. TinySEC establishes authentication mechanisms (with the use of MAC key), encryption of information and protection against duplication of information.

For authentication and encryption TinySEC uses a building cipher block channels with CBC-MAC to create and verify the key MAC. This method is particularly suitable for sensor networks because it does not require a large memory.

Outside these two protocols, the scientific community has proposed other solutions to secure a WSN. (Deng 2005) give a solution to prevent the capture of information in wireless sensor networks by the data partitioning. The aim is to divide the information into several parts. For example a message is divided into 3 packets which are going to follow 3 different routes. If someone want to read the information, he needs to have all the parts.

Another solution proposed by (Zhu 2004), (Ganeriwal 2004) and also by (Oleshchuk 2007), is to use the mechanisms of trust and reputation that can be found in peer to peer networks, community networks or even market websites like Ebay . In this kind of network as in wireless sensor networks, it is hard, to know what node can be a malicious node, because the number of nodes is large. To detect and protect the integrity of the network, each node of the network will monitor its neighboring nodes and their actions over time. Depending on actions taken by its neighboring nodes, a node will increase a level of trust of these nodes, based on its reputation. When a node does not carry out a request, its level of trust will fall. If this node always sends data correctly, its level of trust will increase. With the help of these levels of trust, a node will then choose the most secure route for sending data. Instead of going through the fastest route (number of hops or geographical distance), the node will choose to send its data via nodes with the highest level of trust (the safest route).

6 Presentation of Secured Telemedicine Software

We will present the Servastic project. The aim of this software is to provide practitioners several means of assistance to take the right decisions on complex medical cases. Two tools are available:

1. a module providing image treatment functions: we can send a sequence of images to be analyzed, these functions will return results for a diagnosis assistance. (figure 7) (Covalia 2009).
2. a module providing communication tools to allow a doctor to interact with an expert on complex cases.

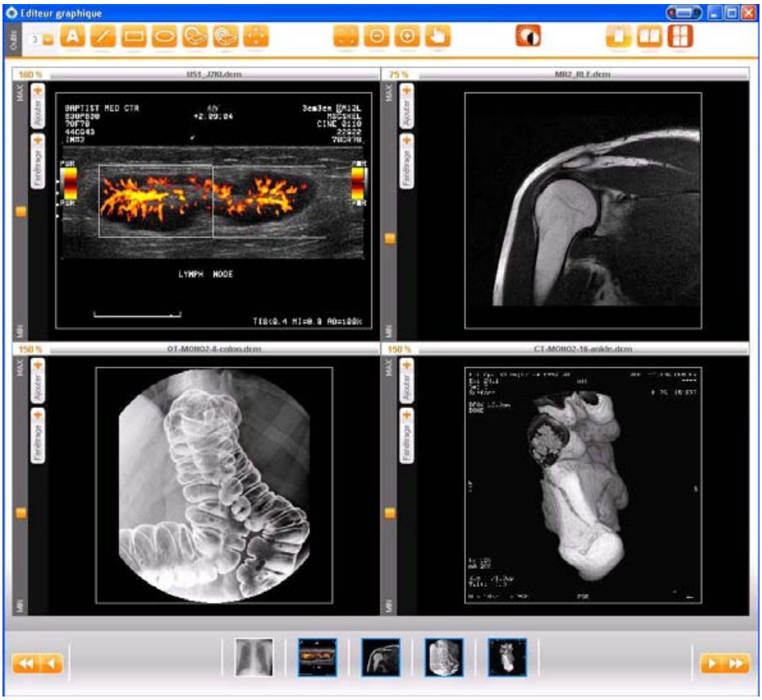


Fig. 7. Secured Telemedicine Software

One of the essential keys of medical imaging evolution is the transition from a qualitative analysis performed by a specialist looking at images, to a quantitative analysis providing precise statistics on these images. This kind of technique is very useful for neuro-degenerative diseases (Alzheimer). At an early stage, anatomic changes not visible by a human can be detected by a computer with a mathematic comparison with reference images (quantitative analysis). This allows the practitioners to save time on diseases diagnosis.

The architecture of the service is presented on figure 8. The aim is to provide centralized computational services on a server. These services have to be accessible from anywhere to allowed users. A main issue is to allow these users to handle complex calculations without having particular software on their computer.

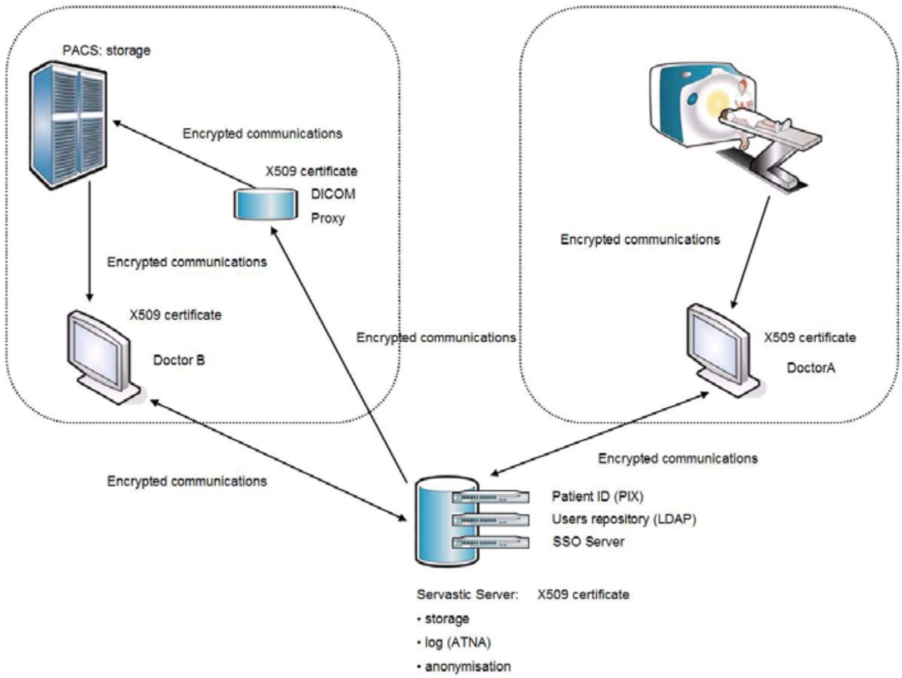


Fig. 8. Software Functionalities

In this case, we use several security technologies. We have a LDAP repository and a SSO server to authenticate allowed users. Each connection and calculation is logged on the server. Encryption and X509 certificate are used to ensure a secured communication (SSL connections), and an anonymization stage is performed on DICOM files (patient ID are removed from the DICOM file, we can keep a mapping with he PIX service if we wish to reuse the data).

The second module allows practitioners to interact on medical data using the network. In this case, we also use authentication, encryption of communication and stored data, anonymization using the PIX service. We have to use a node authentication when a practitioner wishes to insert data sent by a colleague in his own PACS. It can be useful if the expert wants to use its usual interpretation station to manipulate images; in this case the images have to be put in the local PACS. As previously stated, a PACS accepts only communications with authenticated nodes. This operation is quite heavy, since we have to work with hospitals IT services to update the list of authorized nodes. We use a DICOM proxy to make it possible to send medical data from one site to the PACS of another site. For this kind of operation, the link with the patient ID service is essential. Indeed, a patient with an ID "A" in one site can have an ID "B" on the other site. It could be worse if a patient with an ID "A" already exists on the second site. We see clearly here that we have to meet the requirements of patient identification to provide a secured environment.

7 Conclusion

In this chapter we defined telemedicine and the different areas affected. These types of remote medical applications need a management of information security. As the population of the world is growing quickly, more and more physical space will become necessary to store medical data. In addition, patient data are confidential and should not be open to the public. Several services are required to secure the storage of data: file access control, file anonymization, file encryption and secured data access. These security features are now available with several different technologies. Medical field is in advance in standardization for data exchanges (DICOM...). Recommendations begin to appear for security management, with IHE or national security referentials.

The aim of the next years will be to guarantee a high security without making a complex use, deployment and interoperability of telemedicine applications. Security of data transfer and data storage in telemedecine's applications is very important for the development of telemedecine's applications to be aware of the risk that information could be altered if bad solutions are used to secure data. A telemedecine's application must have a strong policy to ensure that information of a patient could not be read, modify or delete.

Acknowledgement

The authors thank:

– the European Community (by the projects TeNeCi, and Decopreme), which allowed a part of the financing of this work,
– as well as the French ministry of Research and the French ministry of industry which allowed financial aids for the creation of the new startup Covalia Interactive.

Resources

A Key Books

Oleg S. Pianykh, Digital Imaging and Communications in Medicine standard (DICOM). *Springer (2008).*

B.S Chowdhry and Faisal Abro, Telemedecine Modernization and Expansion of Health Care Systems. *Mehran Info Tech Consultants (2003).*

B Key Survey/Review Articles

Otto Bernecker, Biometrics security: An end user perspective. Giesecke and Devrient GmbH, Germany, *Information security technical report 11 (2006)*, 111-118.

Butler Lampson, Martin Abadi, Michael Burrows and Edward Wobber, Authentication in Distributed Systems:Theory and Practice. *ACM Transactions on Computer Systems, Vol. 10, No. 4, (November 1992)*, 265–310.

David Martins and Herve Guyennet, Etat de l'art - Securite dans les reseaux de capteurs sans fil. *Conference SAR-SSI (2008).*

Yih-Chun Hu and Adrian Perrig, A Survey of Secure Wireless Ad Hoc Routing. *IEEE Security and Privacy (2004), 28–39.*

C Key Journals

Journal of Methods of Information in Medicine *(Schattauer).*

International Journal of Telemedicine and Applications *(Hindawi Publishing Corporation).*

International Journal of Telemedicine and e-Health *(Mary Ann Liebert, Inc. publishers).*

IEEE Security and Privacy *(IEEE).*

IEEE Medicine and Biology Magazine *(IEEE).*

D Key International Conferences/Workshops

CSCWD - International Conference on Computer Supported Cooperative Work in Design.

DFMA - International Conference on Distributed Frameworks for Multimedia Applications.

E-MEDISYS - E-Medical Systems International Conference.

MMVR - Medicine Meets Virtual Reality Conference.

SECURECOMM - International Conference on Security and Privacy in Communication Networks.

References

[(Abendroth 2006)]Abendroth, J., Seigneur, J.-M.: Deploying the trusted platform module (TPM) to increased fairness and trust in P2P file sharing servents. In: ISP 2006: Proceedings of the 5th WSEAS International Conference on Information Security and Privacy, pp. 202–207 (2006)

[(AuthenticationWorld)]AuthenticationWorld.com: The business of authentication, `http://www.authenticationworld.com/`

[(Bernecker 2006)]Bernecker, O.: Biometrics security: An end user perspective. Giesecke and Devrient GmbH, Germany. Information security technical report 11, 111–118 (2006)

[(Čapkun 2003)]Čapkun, S., Buttyán, L., Hubaux, J.-P.: SECTOR: secure tracking of node encounters in multi-hop wireless networks. In: SASN 2003: Proceedings of the 1st ACM workshop on Security of ad hoc and sensor networks, pp. 21–32 (2003)

[(Deng 2005)]Deng, J., Han, R., Mishra, S.: Countermeasures Against Traffic Analysis Attacks in Wireless Sensor Networks. In: SECURECOMM 2005: Proceedings of the First International Conference on Security and Privacy for Emerging Areas in Communications Networks, pp. 113–126 (2005)

[(Elmarzouqi 2007)]Elmarzouqi, N., Garcia, E., Lapayre, J.-C., Humbert, P.: Use of ACCM Model for teledermatology implementation. In: DFMA 2007, 3rd int. Conference on Distributed Framework for Multimedia Applications, Paris, France, July 2007, pp. 84–89 (2007)

[(Ganeriwal 2004)]Ganeriwal, S., Srivastava, M.B.: Reputation-based framework for high integrity sensor networks. SASN, 66–77 (2004)

[(Garcia 2005)]Garcia, E., Guyennet, H., Lapayre, J.-C., Moulin, T.: Adaptive Tele-application for Remote Neurology Diagnosis. International Journal of Telemedicine and e-Health 11(6), 692–701 (2005)

[(Gouaux 2002)]Gouaux, F.: Ambient Intelligence and Pervasive Systems for the Monitoring of Citizens at Cardiac Risk: New Solutions from The EPI-MEDICS Project. Computers in Cardiology 29, 289–292 (2002)

[(Halperin 2008)]Halperin, D., Heydt-Benjamin, T.S., Ransford, B., Clark, S.S., Defend, B., Morgan, W., Fu, K., Kohno, T., Maisel, W.H.: Pacemakers and Implantable Cardiac Defibrillators: Software Radio Attacks and Zero-Power Defenses. In: SP 2008: Proceedings of the 2008 IEEE Symposium on Security and Privacy, pp. 129–142 (2008)

[(Hu 2003)]Hu, Y., Perrig, A., Johnson, D.: Packet Leashes: A Defense against Wormhole Attacks in Wireless Ad Hoc Networks. In: INFOCOM (2003)

[(Hu 2004)]Hu, Y.-C., Perrig, A.: A Survey of Secure Wireless Ad Hoc Routing. IEEE Security and Privacy, 28–39 (2004)

[(Hu 2006)]Hu, Y.-C., Perrig, A., Johnson, D.B.: Wormhole attacks in wireless networks. IEEE Journal on Selected Areas in Communications, 370–380 (2006)

[(Karlof 2003)]Karlof, C., Wagner, D.: Secure routing in wireless sensor networks: attacks and countermeasures. Ad Hoc Networks, 293–315 (2003)

[(Karlof 2004)]Karlof, C., Sastry, N., Wagner, D.: TinySec: a link layer security architecture for wireless sensor networks. SenSys, ACM, 162–175 (2004)

[(Lamport 1981)]Lamport, L.: Password Authentication with Insecure Communication. Communications of the ACM 24(11) (November 1981)

[(Lampson 1992)]Lampson, B., Abadi, M., Burrows, M., Wobber, E.: Authentication in Distributed Systems: Theory and Practice. ACM Transactions on Computer Systems 10(4), 265–310 (1992)

[(Le Mer 2004)]Le Mer, P., Soler, L., Pavy, D., Bernard, A., Moreau, J., Mutter, D., Marescaux, J.: Argonaute 3D: A real-time cooperative medical planning software on DSL network. In: MMVR12, 12th Annual Medicine Meets Virtual Reality Conference, Newport Beach, California (January 2004)

[(MacDonald 2008)]MacDonald, J.A.: Authentication & key agreement for off-portal mobile applications. Information security technical report 13, 127–135 (2008)

[(Mailhes 2003)]Mailhes, C., Castanié, F., Henrion, S., Lareng, L., Alonso, A.: The URSAFE telemedicine project: improving health care of the elderly. In: MIE, Saint Malo, France, European Federation for Medical Informatics (mai 2003)

[(Malan 2004)]Malan, D., Fulford-Jones, T., Welsh, M., Moulton, S.: CodeBlue: An Ad Hoc Sensor Network Infrastructure for Emergency Medical Care. In: International Workshop on Wearable and Implantable Body Sensor Networks (2004)

[(Marescaux 2001)]Marescaux, J., Leroy, J., Gagner, M., Rubino, F., Mutter, D., Vix, M., Butner, S.E., Smith, M.K.: Transatlantic Robot-Assisted Telesurgery. Nature 413, 379–380 (2001)

[(Marescaux 2002)]Marescaux, J., Leroy, J., Rubino, F., Vix, M., Simone, M., Mutter, D.: Transcontinental Robot Assisted Remote Telesurgery: Feasibility and Potential Applications. Ann. Surg. 235, 487–492 (2002)

[(Martins 2008)]Martins, D., Guyennet, H.: Etat de l'art - Securite dans les reseaux de capteurs sans fil. In: Conference SAR-SSI (2008)

[(Montagnat 2008)]Montagnat, J., et al.: A Secure Grid Medical Data Manager Interfaced to the gLite Middleware. Journal of Grid Computing 1(6), 45–59 (2008)

[(Needham 1978)]Needham, R.M., Schroeder, M.D.: Using Encryption for Authentication in Large Networks of Computers. Xerox Palo Alto Research, Communications the ACM 21(12) (December 1978)

[(Newsome 2004)]Newsome, J., Shi, E., Song, D., Perrig, A.: The Sybil attack in sensor networks: analysis & defenses. In: Third International Symposium on Information Processing in Sensor Networks, IPSN 2004, pp. 259–268 (2004)

[(Oleshchuk 2007)]Oleshchuk, V., Zadorozhny, V.: Trust-Aware Query Processing in Data Intensive Sensor Networks. In: SENSORCOMM 2007: Proceedings of the 2007 International Conference on Sensor Technologies and Applications, pp. 176–180 (2007)

[(OTELO 2004)](OTELO 2004) OTELO clinical and technical validation (2004). Reference D35_WP7_2004_10_1_A (2004)

[(Perrig 2002)]Perrig, A., Szewczyk, R., Tygar, J.D., Wen, V., Culler, D.E.: SPINS: Security Protocols for Sensor Networks. Wireless Networks, 521–534 (2002)

[(Pianykh 2008)]Pianykh, O.S.: Digital Imaging and Communications in Medicine (DICOM), 377 pages. Springer, Heidelberg (2008)

[(Piotrowski 2006)]Piotrowski, K., Langendoerfer, P., Peter, S.: How public key cryptography influences wireless sensor node lifetime. In: SASN 2006: Proceedings of the fourth ACM workshop on Security of ad hoc and sensor networks, pp. 169–176 (2006)

[(Ruud 2001)]Bolle, R.M., Connell, J.H., Ratha, N.K.: Biometric perils and patches. Pattern Recognition 35, 2727–2738 (2002)

[(Stajano 1999)]Stajano, F., Anderson, R.J.: The Resurrecting Duckling: Security Issues for Ad-hoc Wireless Networks. In: Security Protocols Workshop, pp. 172–194 (1999)

[(Seitz 2005)]Seitz, L., Pierson, J.-M., Brunie, L.: Encrypted storage of medical data on a grid. Journal of Methods of Information in Medicine, 198–202 (January 2005)

[(Seung 2001)]Yi, S., Naldurg, P., Kravets, R.: Security-aware ad hoc routing for wireless networks. In: MobiHoc 2001: Proceedings of the 2nd ACM international symposium on Mobile ad hoc networking & computing, pp. 299–302 (2001)

[(Schwiderski-Grosche 2006)]Schwiderski-Grosche, S.: Security: An end user perspective. Information Security Group, Royal Holloway, University of London, Egham, Surrey TW200EX, United Kingdom. Information securitytechnical report 11, 109–110 (2006)

[(Tuyls 2004)]Tuyls, P., Goseling, J.: Capacity and Examples of Template-Protecting Biometric Authentication Systems. In: Maltoni, D., Jain, A.K. (eds.) BioAW 2004. LNCS, vol. 3087, pp. 158–170. Springer, Heidelberg (2004)

[(Vieyres 2003)]Vieyres, P., Poisson, G., Courréges, F., Merigeaux et, O., Arbeille, P.: The TERESA project: from space research to ground tele-echography. Industrial robot: an international journal 30(1), 77–82 (2003)

[(Worth 2003)]Jonathan, L., Worth, M.D., Theodore, A., Stern, M.D.: Benefits of an Outpatient Psychiatric TeleConsultation Unit: Results of a 1-Year Pilot. Primary Care Companion J. Clin. Psychiatry 5, 80–84 (2003)

[(Zhang 2005)]Zhang, Z., Zhang, R., Yao, J., Pu, G.: Medical data on demand with WebmiaEngineering. IEEE Medicine and Biology Magazine 24(3), 117–122 (2005)

[(Zhu 2004)]Zhu, H., Bao, F., Deng, R.H., Kim, K.: Computing of trust in wireless networks. In: Proceedings of 60th IEEE Vehicular Technology Conference, Los Angeles, California (2004)

Web and Software

[(Covalia 2009)](Covalia 2009) Covalia Interactive is the Covotem producer
http://www.covalia.com/

[(ISfTeH 2009)](ISfTeH 2009) International Society for Telemedicine and eHealth
http://www.isft.net/

[(Mlabs 2009)](Mlabs 2009) Multimedia Conferencing System producer
http://www.mlabs.com/

[(SET 2009)](SET 2009) European Society of Telemedecine
http://www.societetelemed.eu/

[(TinyOS 2009)](TinyOS 2009)TinyOS: an open-source operating system designed for wireless embedded sensor networks.
http://www.tinyos.net/

Chapter 10

Audio Content Analysis in the Urban Area Telemonitoring System

Paweł Żwan, Przemysław Sobala, Piotr Szczuko, and Andrzej Czyzewski

Multimedia Systems Department, Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, Poland

1 Introduction

Monitoring systems of urban areas comprise a large number of video cameras, occasionally with an inclusion of other sensors, such as: RFID readers, IR motion detectors, access cards and others. Dangerous events may also be accompanied with a sound cues, that, if analyzed by the monitoring system, may reveal more information than only video analysis. Therefore utilization of cameras with microphones attached is a current trend in evolution of monitoring systems. Sound analysis may provide an effective detection of screams, calling for help, breaking of a glass, explosions, gun shots, etc. Methods are developed for estimation of direction of sound arrival, therefore allowing to point a moving video monitoring camera on the sound source. Combining both techniques, audio analysis and video analysis in single monitoring system results in a great increase of the threat detection efficiency.

The chapter presents various aspects of audio signals processing dedicated to detection and monitoring of threats in urban areas. The audio analysis is described in Sec. 2. It comprises detection of sounds and classification of them into 4 groups: broken glass, explosion, scream, and environmental sounds non-related to threats. Classification algorithm and results are presented in Sec. 2.1. Once in an audio stream an important event is detected, it is crucial to estimate localization of the sound source. Sec. 2.2 describes techniques for the localization of sound source utilizing set of microphones. Outcome data of localization module can be utilized to drive moving PTZ cameras correctly, as is discussed in Sec. 3.

2 Audio Analysis for Detection of Threats

It is assumed that beside the video stream the other medium analyzed by the monitoring system is audio stream captured by the microphones of the monitoring station. The audio information complements the visual content of a multimedia data in a sense of the parametric description. The mechanisms for detection and classification of

G.A. Tsihrintzis and L.C. Jain (Eds.): Multimedia Services in Intel. Environ., SIST 3, pp. 227–239.
springerlink.com

sounds were developed. Location of sound source can be estimated utilizing sound analysis coming from a number of microphones.

2.1 Automatic Classification of Sounds for Security Purposes

The algorithm allowing automatic recognition of sound events is presented in this chapter and discussed. Although the methods of the parameterization of sound are described in a rich literature (e.g. related to MPEG-7 standard) [1], the automatic analysis of the audio data stream in this case demands a particular effort. It faces many difficulties, of which a non-stationary character of the background noise plays the most important role. In the case of the outdoor audio monitoring the noise usually reaches a significant level, so that its RMS value can sometimes be even higher than the RMS level of the interesting audio event. Moreover, in the environment there may propagate loud sounds which ought not be erroneously recognized by the system as dangerous events (e.g. the sounds of passing cars or the sounds of the atmospherics such as a strong rain, thunders etc). In this aspect a classical de-noising spectral subtraction methods cannot be directly used, but they need to be modified in order to match this special application field. The developed algorithm is based on a spectrum analysis of succeeding audio frames and on a continuous computing of a sonogram. The background noise is monitored by averaging the spectrogram in a selected period of time and then memorized. The differential signal is formed by a subtraction of the analyzed signal and the noise approximation with a weight which value is optimized. The denoised sound events are then parameterized and automatically interpreted.

The system is designed to work directly on audio streams, however for evaluation it was trained and tested on sound files. For that purpose a database of sound samples needed to be formed. It consists of examples of different audio events. For the group of the sounds which can be related to a danger for citizens three “danger classes” have been defined: a broken glass (C1), an explosion (gunshot) (C2) and a scream (C3). Fourth category of sounds C4 represents all other sounds not related to any danger (various environmental sounds). Since the risk of false alarms needs to be minimized it is crucial to not relate wrongly environmental sounds of the class C4 to the classes C1–C3.

The sounds were recorded in a noiseless environment and the number of gathered sound samples was 40 for the category C1, 36 for the category C2 and 28 for the category C3. The Category of “not danger” environmental sounds (C4) have been represented by the streams of recorded city traffic sounds. Since it was important that the algorithm for different noise types can be tested, thus the ‘background tracks’ were recorded as 10 noise samples coming from different parts of a city. In the experimental part during the the algorithm testing, recognized sounds were simulated by mixing up the sounds from classes C1–C3 with background samples of different types. The Class C4 has been represented by the background track itself. In the summation the levels of RMS of the background track and the sound events (C1–C3) were equal. It simulated a strong noise environment.

The main modules of the recognition system are presented in Fig. 1. For a given audio event a number i of audio sequences $x_i(n)$ (audio streams) is fed to the modified spectral subtraction algorithm (MSSA) and the output sequence $r_i(n)$ is analyzed, and finally their set of parameters $\{P\}$ is extracted. Since the analyzed audio signal must be divided into frames, n in the Fig. 1 represents the number of the sample in an analyzed frame. The similar approach has been used by the authors in other works related to the automatic sound recognition [2][3], proved to be successful in some other multimedia applications. The methodology seems similar as in case of different sound recognition systems, but in this particular application it is crucial to propose some proper parameterization methods. Each of the blocks of the scheme in Fig. 1 is presented and discussed in the next parts of the chapter.

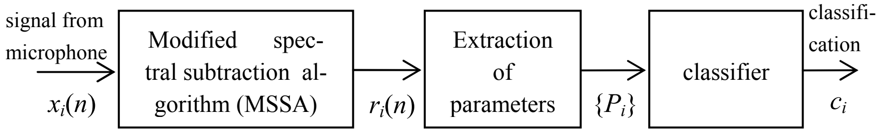


Fig. 1. Modules of the automatic sound classification algorithm

The proposed system works on the basis of the spectral subtraction algorithm [4][5][6] which has been modified in order to match the task of the automatic sound recognition in a strong non-stationary noise. It is based on the averaging of the sound spectrogram in the time domain for a given period of time in order to approximate spectral characteristics of the noise. The modification of the classical approach is that the averaged noise spectral approximation result is then subtracted from the actual spectrum of the analyzed frame $x_i(n)$ with a given weight α. Assuming a short character of the sound events and a proportionally long averaging time, the system remains stable and has the ability of working in very noisy environments. An adequate parameterization of audio events, which is presented later, allows for classifying of those events into one of the predefined categories. The schema of the elaborated algorithm is presented in Fig. 2.

The signal is divided into frames of 4096 samples – $x_i(n)$ and than the Fast Fourier Transform $f_i(n)$ is calculated. Next, a logarithmic power spectrum $l_i(n)$ is obtained and several subsequent analyses $l_i(n)$ form a spectrogram. In the next step an averaged audio spectrum $Af_i(n)$ from k previous samples is created according to Eq. (1) and the spectral subtraction (2) is performed:

$$Af_i(n)=(Af_i(n)\bullet k+f_i(n))\bullet(k+1) \quad (1)$$

$$r_i(n)=a_i(n)-\alpha\bullet Af_i(n) \quad (2)$$

where n is a sample number, α and k are analysis parameters.

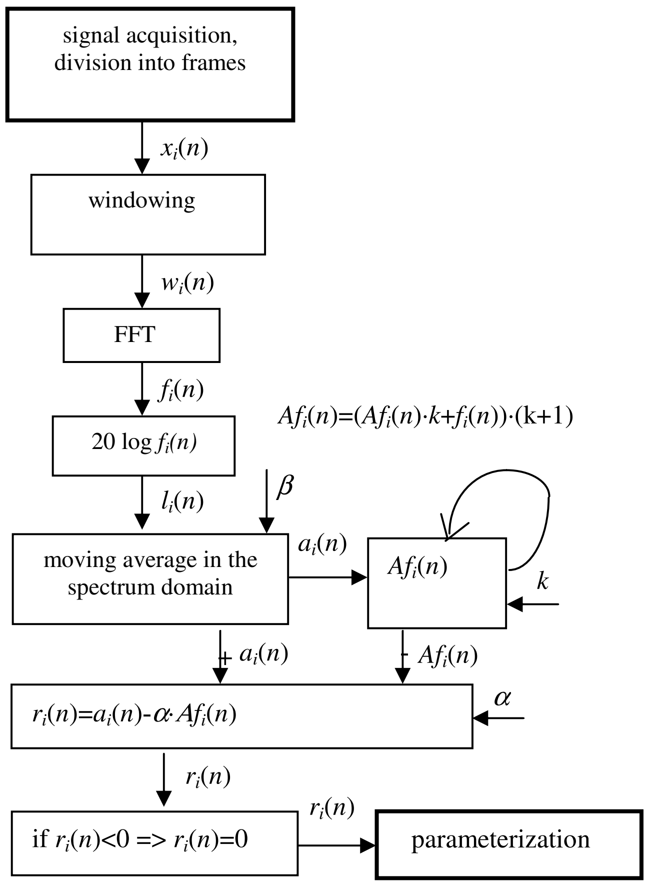


Fig. 2. Modified algorithm of a spectral subtraction

The subtraction equation has an additional parameter α whose value must be optimized. The last operation is the thresholding of the resulting sequence $r_i(n)$ by forcing to zero the values of those elements with a level less than 0. The parameters of the algorithm have been optimized. The methodology was described in literature [7]. The optimization was based on the analysis of the automatic recognition accuracy for the optimal values of k, α and β are α=0.4, β=16 and k=50. Those parameters can be interpreted directly. The frame of 16 samples is used for averaging the spectrum, 50 previous analyses are used to obtain an averaged spectrum of the noise and then the spectrum is being subtracted from the power spectrum of the sequence $x_i(n)$ with a weight of 0.4.

The set of sequences $r_i(n)$ can be written as a matrix of dimensions $i\times n$ and this matrix is put forward for the proper parameterization. Since each of sequences $x_i(n)$

comes from the FFT analysis, the domain of $x_i(n)$ can be related to the domain of frequency with a simple relation (Eq. 3):

$$f=n\cdot(f_s/N) \tag{3}$$

where f and n are the frequency and the sample number, f_s is the sampling frequency and N the length of the analyzed frame.

In the first step of the parameterization each one of the sequences $r_i(n)$ is divided into three sub-sequences related to the three frequency bands a) f=<0, 500Hz>, b) f=(500Hz, 2700Hz> c) f=(2700 Hz, 8000Hz). The boundary values 500Hz, 2700Hz and 8000Hz result from the analysis of the acoustic signals, e.g. around the frequency 2700 Hz the energy of harmonics of the signal of the human-originated scream includes values -20dB less than the value of the its maximum harmonic. According to Eq. 3 the boundary frequency values f=500Hz, f=2700Hz and f=8000Hz are related to samples number: n=46, n=251 and n=744 (sample length N=4096, sampling frequency f_s=44100). The values y_i^m are created by summing up the sequence $r_i(n)$ as it is presented in Eq. 4.

$$y_i^m = \sum_{n=N_1(m)}^{N_2(m)} r_i(n) \tag{4}$$

where i is time index, m=1..3, $N_1(m)$={0,47,252), $N_2(m)$={46,251,744}.

An event is recognized as non-typical for a given acoustical background if the y_i^m for almost two values m exceeds the threshold levels (set experimentally). The recognized event is then additionally parameterized. 20 subsequent frames are taken to the analysis and parameterized in relation to the time envelope of the y_i^m (i=1..20). The time and frequency analysis of the sequence of y_i^m allows the recognition of such events as a broken window, a gunshot or an explosion. 20 subsequent values of y_i^m are denoted in the matrix $Y^m(n)$ where n=1…20, m=1,2,3. Each one of those 3 sequences ($Y1(n)$, $Y2(n)$, $Y3(n)$) is parameterized by using of the following parameters:

- $p1_m=max5(Y^m(n))$, where '*max5*' is the medium value of 5 greatest elements of the sequence $Y^m(n)$,
- $p2_m=\text{count}(Y^m(n)>Y_{thr}^m)$, where '*count*' is the number of the elements of the sequence $Y^m(n)$ with the values greater than the threshold value Y_{thr}^m, where the threshold value Y_{thr}^m is the function of the number of the sub-band m. $Y_{thr}^m=\text{f}(m)$.

The set of the defined parameters $\{p1_1, p1_2, p1_3, p2_1, p2_2, p2_3\}$ allows for the definition of the next descriptors, defined in numerous practical experiments.

Parameter $p3$ is the centre gravity value of the parameters $p1_m$, m=1...3.

$$p3 = \frac{\sum_{m=1}^{3}(p1_m \cdot m)}{\sum_{m=1}^{3}(p1_m)} \tag{5}$$

The parameter $p4$ is the arithmetical mean of the parameters $p1_2$ i $p1_3$.

$$p4 = \frac{p1_2 + p1_3}{2} \tag{6}$$

The parameter $p5$ is the centre gravity value of the parameters $p2_m$, m=1...3.

$$p5 = \frac{\sum_{m=1}^{3}(p2_m \cdot m)}{\sum_{m=1}^{3}(p2_m)} \tag{7}$$

The parameter $p6$ is the arithmetical mean of the parameters $p2_m$, m=1…3.

$$p6 = \frac{\sum_{m=1}^{3} P1_m}{3} \tag{8}$$

The detection of the scream in the acoustic signal demands an additional parameterization related to the acoustic properties of the scream signal, which occupies the characteristic frequency band and possesses a significant level of harmonics.

$$k = \left(\log \frac{\sum_{30}^{600} r_i(n)}{\sum_{1}^{29} r_i(n) + \sum_{601}^{1024} r_i(n)}\right) + \left(\max\big(s_i(n)\big) - \frac{\sum_{30}^{600} r_i(n)}{600-30}\right) \tag{9}$$

where $s_i(n)$ is the sequence $ri(n)$ created by 570 samples between indexes 30 and 600.

The first segment of the equation (9) is related to the calculation of the ratio of the energy of $r_i(n)$ in the frequency band n=<30,600> (f=<323Hz,6460Hz>) to the energy in the remaining part of $r_i(n)$. The second segment is the difference between maximum of $s_i(n)$ and the mean value of the samples in the $s_i(n)$. Both $r_i(n)$ and $s_i(n)$ result from the logarithmic analysis, thus additional logarithmic function in the second segment of Eq. 9 is not needed. The value of k is also calculated in subsequent 20 time frames forming a vector $K(n)$, n=1…20. The value of the parameter $p7$ is defined as the medium value of the 5 six maximum elements in the sequence $K(n)$ (analogically as, was defined for the parameter $p1_m$).

In order to complement the set of parameters, the spectral moment of the sequence $r_i(n)$ is also defined (Eq. 10). The center of gravity of the sequence $r_i(n)$ allows for a better determination of the scream signal and complement the information given by a parameter $p7$.

$$p8 = \frac{\sum_{n=1}^{744} r_i(n) \cdot n}{\sum_{n=1}^{744} r_i(n)} \tag{10}$$

The higher limit of the sum was set to n=744. This value is related to the frequency 8kHz. Above that frequency no significant features are presented in the spectrum for any of recognized classes of the sounds.

The extracted parameters have been calculated for all sounds from the database (classes C1-C4). The database consists of 104 sound samples. Since the functionality of the system has to be verified in normal circumstances, the recognized samples has been added to randomly chosen noise samples from 10 recorded environments. For each of the sound samples three recognized objects have been formed (with 3 different noise samples) which has formed 324 objects . Basic on the objects the recognition system have been evaluated.

The parameters can be presented by using graphs with axes related to chosen pairs of parameters. The examples of such analyses are presented in Fig. 3.

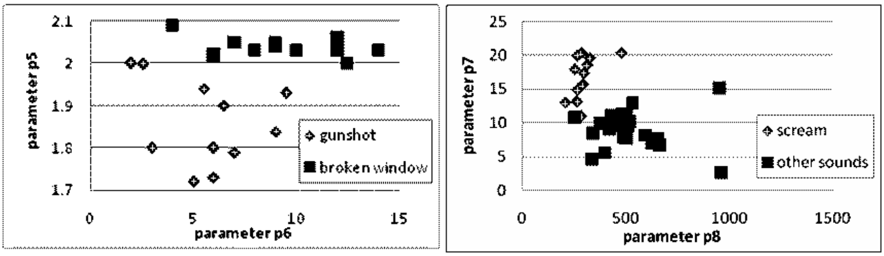


Fig. 3. Distribution of the values of some chosen parameters for examples of analyzed sounds in 2D parameters plots

Since there exists a significant difference in the values of some of the parameters for different classes, it was decided to use a linear decision system – the decision tree. The methodology is as follows: values of each of the parameters have been statistically checked using the Fisher statistics (FS) values for each pair of the classes [8]. The parameters with a significant value of FS for a given pair of classes have been considered as important. This analysis is described in literature [7].

The decision is taken by the following algorithm. First, the parameter $p4$ decides whether the sound can be classified to one of the following groups: 'scream'/'other sound' and 'broken glass'/'gunshot'. Secondly, the recognition within the first group is performed by using the parameters $p7$ and $p8$. The acoustical signals are classified as the scream if parameter $p7>12$ and $p8<500$. The recognition between broken glass and gunshot sounds is performed by parameters $p5$, $p6$ and $p1_1$. If the values of ($4<p6<6$ and $p5<2$) or ($p5=2$ and $p6<6$) the decision cannot be taken.

For all 324 recognized objects the average recognition accuracy was 90%. The accuracy was the highest for the category C3 (scream) – 93.5% and the worst for the case of C2 (gun-shot): 85%. In order to check the frequency of the false alarms it has been tested only for ambient sounds (10 recorded environmental sounds). For the 10 minutes of the test signal the false alarm was generated only 2 times (a loud car horn was recognized as a scream). Future experiments will be focused on algorithm optimization in field test, when the accuracy of the automatic recognition can be further improved and the frequency of false alarms diminished.

The system's reaction to detected dangerous situation is discussed in Sec 3.

2.2 Localization of Sound Source

Assuming that in audio stream an important event was already detected, it is crucial to estimate localization of its sound source. That data can be utilized to correctly drive moving PTZ cameras, that should be pointed to the source, so that event video recording should be started. The sound localization techniques are outlined below.

Based on localization method present in human neuro-auditory system, the methods were developed. The basic fact is that for a number of sensors positioned in various points in the space, the arrival times and levels of received sounds mayvary, therefore it is possible to estimate the direction of the sound source. Humans' "two-receiver setup" is presented in Fig. 4. It is always assumed that the source is far and therefore sound waves are propagating in parallel. A person can localize the source basing on two clues:

- ILD – Interaural Level Difference or IID – Interaural Intensity Difference,
- ITD – Interaural Time Difference, visualized in Fig 4, described by Eq. (11).

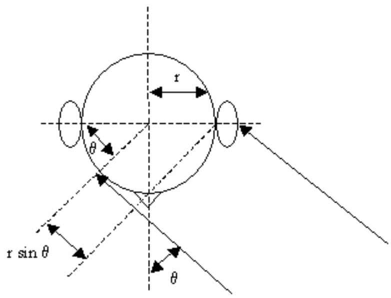


Fig. 4. Graphical representation of time difference of arrival of sound wave to two human ears

$$ITD = \frac{r}{c}(\theta + \sin\theta), \qquad -90° \le \theta \le 90° \tag{11}$$

ITD value differs around the value of $0\,\mu s$, for sounds arriving exactly from the front or back, up to $700\,\mu s$ for directions $\theta = \pm 90°$. For low frequencies, when wavelength is comparable to the head size, ITD localization is less effective.

For microphones in the ULA configuration (Uniform Linear Array), the front-back error occurs, because there is no clue to correct differentiation of these two positions, as time delays for the source in the front are the same as for source positioned in the back (Fig. 5). More complex configurations of microphones are developed for dealing with described problem, nevertheless in the further discussion the ULA is taken into consideration, representing a basic, and a very clear example of direction of sound arrival techniques.

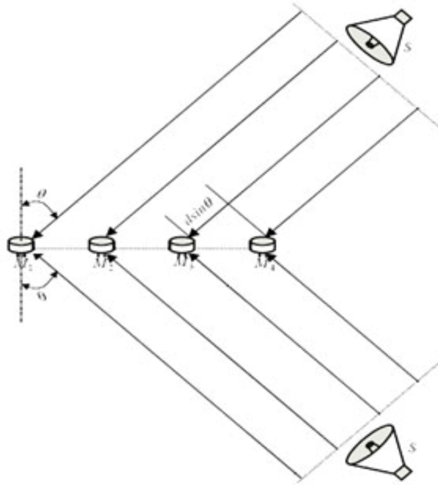


Fig. 5. Front-back localization problem for linear arrays of microphones

The precision of θ calculation based on Eq. (11) depends on the accuracy of estimation of time differences of arrivals (TDOA) for each sensor. The most common method for TDOA calculation is the Generalized Cross-Correlation function – GCC. In that method first a correlation of time-shifted signals is calculated, then a peak value is find, and finally the shift for which the peak occurs is interpreted as a time difference between two signals. The cross-correlation of two signals is calculated as follows:

$$R_{x_1,x_2}(\tau) = \int_{-\infty}^{\infty} W(\omega) X_1(\omega) X_2^*(\omega) e^{j\omega\tau} d\omega \tag{12}$$

where $X_1(\omega)$ and $X_2(\omega)$ are Fourier transforms of $x_1(t)$ and $x_2(t)$, and $W(\omega)$ is weight function. There are five weight functions proposed in the literature [9]:

- Roth Impulse Response,
- Smoothed Coherent Transform (SCOT),
- Phase transform (PHAT),
- Eckart Filter,
- Maximal likelihood function (ML).

For all calculated R coefficients the correct one is selected as a peak value, and final value of time-shift between two signals is obtained [10]:

$$\tau = \arg\max R_{x_1,x_2}(\tau) \tag{13}$$

For multi-microphone setup time-shifts τ are calculated for each pair of sensors, and then, considering spacing between them and orientation in space, a set of θ candidates is obtained originating from each pair. Then utilizing a statistical approach, a single value θ can be estimated, e.g. taken as a median or mean of candidates.

A sample of signals recorded by 6 microphones in ULA with spacing d=0,05m is presented in Fig. 6. A current direction of arrival was $\theta = 90^{\circ}$ – the source was in-line with the microphones.

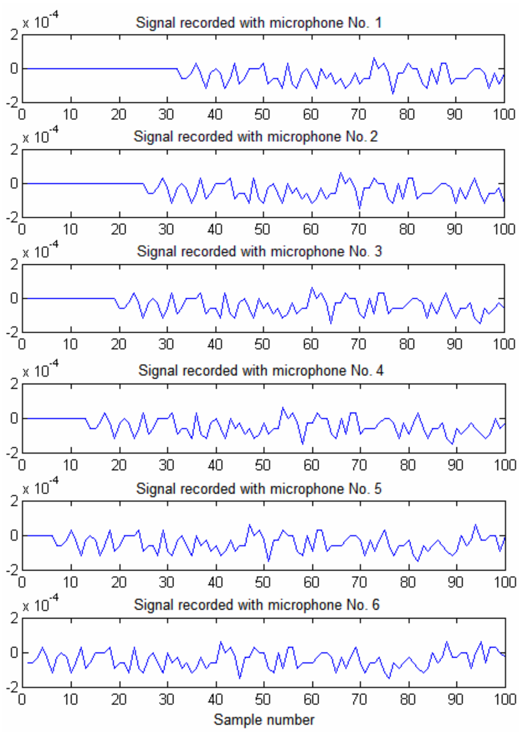


Fig. 6. Signals recorded with 6 microphones in ULA configuration

In conducted experiments various numbers of microphones were used, with different spacing and with variable source localization [11a]. For ULA comprising 8 microphones with spacing 0,05m, utilizing 4 methods for time-shifts estimation and mean/median calculation of result θ following results were obtained (Tab. 1).

The following conclusions can be formulated on the basis of performed experiments:

- the higher is the number of microphones the lesser is the error of localization, but above 8 microphones the further gain of accuracy is minor,
- the larger microphone spacing is the lesser is the localization error,

Table 1. Angle estimation errors depending on used method for ULA with 8 microphones spaced by 0,05m.

	Angle											Mean error	RMS
Method	0	π/8	π/4	π/3	3π/8	π/2	5π/8	2π/3	3π/4	7π/8	π		
I	0,092	0,014	0,005	0,001	0,006	0,000	0,006	0,001	0,005	0,014	0,092	0,021	0,0008
II	0,000	0,014	0,000	0,005	0,011	0,000	0,011	0,005	0,000	0,014	0,000	0,005	0,0000
III	0,064	0,053	0,069	0,058	0,055	0,000	0,056	0,058	0,069	0,053	0,064	0,054	0,0016
IV	0,000	0,053	0,038	0,028	0,040	0,000	0,040	0,028	0,038	0,053	0,000	0,029	0,0006
V	0,525	0,155	0,136	0,105	0,009	0,000	0,009	0,104	0,136	0,155	0,525	0,169	0,0299
VI	0,000	0,053	0,000	0,020	0,016	0,000	0,016	0,020	0,000	0,053	0,000	0,016	0,0003
VII	0,120	0,035	0,051	0,065	0,048	0,000	0,058	0,181	0,024	0,179	0,201	0,088	0,0061
VIII	0,000	0,053	0,027	0,003	0,002	0,000	0,021	0,033	0,027	0,053	0,000	0,020	0,0004

I – CC mean, II – CC median, III – GCC mean, IV – GCC median, V – GCC+PHAT mean, VI – GCC+PHAT median, VII – GCC+ROTH mean, VIII – GCC+ROTH median

- for all methods for angle θ close to 90° the localization error is the smallest and it grows up to 0,5 rad (ca. 30°) while θ decreases to 0° or increases to 180°.
- best methods (Cross Correlation with median and mean θ calculation) provides the accuracy of 0,1 rad (ca. 6°) in the whole range of angles.

In Sec. 3 a utilization of θ value for the detected sound source in the urban area monitoring systems is discussed.

3 Integration of Audio Analysis Module with Monitoring System

Combining techniques described in previous sections: recognition of sound events, estimating sound source localization, and a PTZ (pan-tilt-zoom) positioning of the video camera commonly applied in monitoring systems to observe known point in space, a novel approach for the urban area monitoring system was proposed. It is based on the recognition of audio events related to dangers and on tracking sound sources with a camera.

It was assumed that the monitoring system is equipped not only with cameras but additionally with microphone arrays, e. g. organized as the ULA configuration described in Sec. 2.2. The monitoring station is continuously processing sound signals acquired by the microphones, and performs a detection of sounds related to dangerous events. An audio buffer is employed for keeping current audio data as long as the processing is performed. If important event is detected in a processed data segment, then all audio data gathered in the buffer are stored and transmitted to the center for validating by an operator.

Automatically a localization procedure is activated, sound source is positioned and the information is fed to PTZ camera control module. At the moment when audio event was detected, a recording of audio and video signals is started. The system operators are alarmed about the event, so that they can view the live stream representing the scene or rewind back to the start of the event for a validation of recognition results.

4 Summary

In many cases the sound can make an important factor for detection of dangerous event occurring in a city agglomeration. Therefore, utilizing cameras with microphones attached represents a current trend in the evolution of monitoring systems. The sound analysis may provide an effective detection of screams, calling for help, breaking of a glass, explosions, gun shots, etc. Methods were developed for estimating a direction of sound arrival with an absolute error less than 6°, therefore allowing to point a moving video monitoring camera to the sound source. Combining both techniques, audio analysis and video analysis in a monitoring system may result in an evident increase of threat detection efficiency.

Acknowledgments. The research leading to these results has received funding from the European Community‘s Seventh Framework Program (FP7/2007-2013) under grant agreement N° 218086.

Resources

A Key Books

Cohen L. (1995) Time-Frequency Analysis, Prentice-Hall Inc., New Jersey.

Theodoridis S., Koutroumbas K. (2005) Pattern Recognition, 3d edition, Academic Press.

Manjunath B. S., Salembier P., Sikora T. (2002) Introduction to MPEG-7: Multimedia Content Description Interface, Willey

B Key Survey/Review Articles

Ballas J., Howard H., *Interpreting the Language of Environmental Sounds*, Environment & Behaviour, vo. 19, No. 1, pp. 91-114, Sage Publications, 1987.

Cowling M., Sitte R., *Comparison of Techniques for Environmental Sound Recognition*, Pattern Recognition Letters, Elsevier Science Inc., vol. 24, Iss. 15, pp. 2895-2907, 2003.

Li D., Seithi I.K., *Classification of general audio data for content-based retrieval*, Pattern Recognition Letters, vol. 22, pp. 533-544, 2001.

Valenzise G., Gerosa L., Tagliasacchi M., Antonacci, F., Sarti, A., *Scream and gunshot detection and localization for audio-surveillance systems*, Advanced Video and Signal Based Surveillance, pp. 21-26, 2007.

C Key Journals

Journal of Digital Forensic Practice, Taylor & Francis.

Digital Signal Processing, Academic Press.

Eurasip Journal on Audio, Speech, and Music Processing, Hindawi Publishing Corporation.

D Key International Conferences/Workshops

Hoyt J., Wechsler H., *Detection of Human Speech in Structured Noise*, Proc. ICASSP, vol.2, pp. 237-240, 1994.

Cheng-Yao Chen, Abdallah, A., Wolf, W., *Audiovisual Gunshot Event Recognition*, Systems, Man and Cybernetics IEEE International Conference, Vol. 6, pp. 4807-4812, 2006.

Maher R.C., *Modeling and Signal Processing of Acoustic Gunshot Recordings*, Digital Signal Processing Workshop, pp. 257-261, 2006.

Luigi Gerosa, *Scream and gunshot detection in noisy environments*, 15th European Signal Processing Conference (EUSIPCO-07), Sep. 3-7, Poznan, Poland, 2007.

References

[1] Lindsay, A., Herre, J.: MPEG-7 and MPEG-7 Audio - An Overview, vol. 49(7/8), pp. 589–594 (2001)

[2] Kostek, B.: Perception-Based Data Processing in Acoustics. In: Applications to Music Information Retrieval and Psychophysiology of Hearing. Series on Cognitive Technologies. Springer, Heidelberg (2005)

[3] Kostek, B., Zwan, P., Dziubinski, M.: Musical Sound Parameters Revisited. In: Proc. Music Acoustics Conference, Stockholm, pp. 623–626 (2003)

[4] Kim, W., Kang, S., Ko, H.S.: Spectral subtraction based on phonetic dependency and masking effects, vol. SP(147)(5), pp. 423–427 (2000)

[5] Vizinho, A., Green, P., Cooke, M., Josifovski, L.: Missing data theory, spectral subtraction and signal-to-noise estimation, for robust ASR: An integrated study. In: Proc. Eurospeech 1999, pp. 2407–2410 (1999)

[6] Vaseghi, S.V.: Advanced Digital Signal Processing and Noise Reduction, 2nd edn. Inter-Science, Hoboken (2001)

[7] Zwan, P.: Automatic sound recognition for security purposes. In: Proc. 124th Audio Engineering Society Convention, Amsterdam, The Netherlands (2008)

[8] Sebestyen, G.S.: Decision-making processes in pattern recognition. Macmillan Publishing, Indianapolis (1965)

[9] Knapp, C., Carter, G.: The Generalized Correlation Method for Estimation of Time Delay. IEEE Transactions on Acoustics, Speech and Signal Processing 24(4), 320–327 (1976)

[10] Park, B.C., Ban, K.D., Kwak, K.C., Yoon, H.S.: Sound Source Localization Based on Audio-visual Information for Intelligent Service Robots. In: International Symposium of Information Systems (2007)

[11] Sobala, P.: Methods of audio signal processing for the monitoring systems (in polish). MSc Thesis, Gdańsk University of Technology (2008)

Chapter 11

Video Content Analysis in the Urban Area Telemonitoring System

Piotr Dalka, Grzegorz Szwoch, Piotr Szczuko, and Andrzej Czyzewski

Multimedia Systems Department, Faculty of Electronics,
Telecommunications and Informatics,
Gdańsk University of Technology, Poland

1 Introduction

The task of constant monitoring of video streams from a large number of cameras and reviewing the recordings in order to find a specified event requires a considerable amount of time and effort from the system operators and it is prone to errors. A solution to this problem is an automatic system for constant analysis of camera images being able to raise an alarm if a predefined event is detected.

The chapter presents various aspects of video signals processing dedicated to detection and monitoring of threats in urban areas. First the video analysis methods are presented, then recognition algorithms are introduced for detection of important events, and finally obtained results are discussed.

The section 2. is dedicated to basic video analysis algorithms aimed at detection, tracking and classification of moving object appearing in a camera field of view. First in section 2.1 moving object detection is presented. It employs background modelling and subtraction for determination of non-stationary objects detection. Then the problem of objects shadows is discussed, and finally a stage of image segmentation is presented. In Sec. 2.2 moving object tracking is described. This process entails a necessity of solving many problems regarding object occlusions. An approach utilizing changes in time of object state described by Kalman filter is discussed, which concludes low level processing of visual media.

An outcome of low-level processing serves as a input to higher-level analysis, where it is determined whether they are actually dangerous situations. The high-level video processing is presented in Sec. 3. It starts with object classification (Sec 3.1), and then event recognition is performed (Sec. 3.2). Modern video surveillance systems contain multiple cameras, covering a wide area, some of which are able to pan, tilt and zoom view area (PTZ cameras). A technique for positioning of PTZ cameras is presented in Sec. 3.3, allowing tracking of moving object.

G.A. Tsihrintzis and L.C. Jain (Eds.): Multimedia Services in Intel. Environ., SIST 3, pp. 241–261.
springerlink.com

2 Basic Video Analysis

The main goal of the basic video analysis for the purpose of the intelligent telemonitoring and surveillance system is to detect, track and to classify all moving object appearing in a camera field of view and to calculate various low-level descriptors which are utilized by higher-level modules, such as an event detector, discussed later in Sec. 3.

2.1 Moving Object Detection

Moving object detection and segmentation is an important part of video based applications, including video surveillance. In the latter application, results of detection and segmentation of objects in video streams are required to determine the type of an object and to classify events related to the object.

Most video segmentation algorithms usually employ spatial and/or temporal information to generate binary masks of objects [1][2][3]. The solution presented here utilizes spatial segmentation to detect moving objects in video sequences. The most popular region-based approach is background subtraction [4]. However simple time-averaging of video frames is insufficient for a surveillance system because of limited adapting capabilities.

In our surveillance system, the adaptive background subtraction method modelling pixels as mixtures of Gaussians and using an on-line approximation to update the model was employed [5][6]. This method proved to be useful in many applications, as it is able to cope with illumination changes and to adapt to the background model accordingly to the changes in the scene, e.g. when motionless foreground objects eventually become a part of the background. Furthermore, the background model can be multi-modal, allowing regular changes in the pixel colour. This makes it possible to model such events as trees swinging in the wind or traffic light sequences.

2.1.1 Background Modelling

Background modelling is used to model current background of the scene and to differentiate foreground pixels of moving objects from the background [7][8]. Each pixel in the image is modelled with a mixture of K Gaussian distributions for this purpose,. The probability that a pixel has the value x_t at the time t is given as:

$$p(x_t) = \sum_{i=1}^{K} w_t^{\,i} \eta\left(x_t, \mu_t^{\,i}, \Sigma_t^{\,i}\right) \tag{1}$$

where w_t^i denotes the weight and μ_t^i and Σ_t^i are the mean vector and the covariance matrix of ith distribution at the time t, and η is the normal probability density function:

$$\eta(x_t, \mu, \Sigma) = \frac{1}{(2\pi)^{0.5\cdot D}\sqrt{|\Sigma|}} e^{-0.5\cdot(x_t-\mu)^T \Sigma^{-1}(x_t-\mu)} \tag{2}$$

where D is the number of elements describing pixel colour; for the RGB colour space D is equal to 3.

The number of distributions K is usually a small number between 3 and 5 and is limited by the available computational power.

It is assumed, that each Gaussian distribution represents a different background colour of a pixel. The longer a particular colour is present in the video stream, the higher value of the weight parameter w of the corresponding distribution.

With every new video frame, the parameters w, μ and Σ of distributions for each pixel are updated according to the on-line K-means approximation algorithm. In the first step, distributions are ordered based on the value of the r coefficient given as:

$$r = \frac{w}{\sqrt{|\Sigma|}} \tag{3}$$

where $|\Sigma|$ is the determinant of the covariance matrix Σ. A particular colour of the scene background is usually more often present in the observation data than any colour of foreground objects and as such is characterized by the low variance. Thus a distribution with a higher r value represents the background colour more accurately.

Every new pixel value x_t is checked against existing distributions, starting from the distribution with the highest value of the r coefficient, until the first match is found. The pixel matches the distribution if its colour lies within 2.5 standard deviations of the distribution. If there is no match, a distribution with the lowest r value is replaced with the new one with the current pixel as its mean values, an initially low weight and high variances.

The weight of the first matching distribution is increased, while the weights of other distributions are decreased based on the value of the learning rate parameter α [7]. The higher the α is the faster model adjusts to changes in the scene background (e.g. caused by gradual illumination changes), although moving objects, which remain still for a longer time (e.g. vehicles waiting at traffic lights), would quicker become a part of the background.

If there is a matching distribution, its mean and variance values are tuned according to the current value of the pixel; the speed of converging is determined by the learning rate α [7].

Only the first D distributions of pixel x in time t ordered by the decreasing r coefficient value are used as the background model where D is defined as:

$$D_x^t = \arg\min_d \left(\sum_{i=1}^{d} w_i^{\,t} > T \right) \tag{4}$$

If T is small, then the background model is usually unimodal. If T is higher, the background colour distribution may be multi-modal, which could result in more than one colour being included in the background model. This make possible to model periodic changes in the background, properly. If the current pixel value does not match any of the first D distributions, it is considered as a part of a foreground object.

2.1.2 Shadow Detection

Shadow detection and removal is required for every outdoor video processing application, especially in the field of video surveillance. The shadow of a moving object moves together with an object and as such is detected as a part of the foreground object by a background removal algorithm.

The shadow detection method is based on the idea that while the chromatic component of a shadowed background part is generally unchanged, its brightness is significantly lower [7][9]. Every new pixel recognized as a part of a foreground object during the background subtraction process is checked whether it belongs to a moving shadow. If for at least one of the first D distributions forming the background model the current pixel is darker than the distribution and its colour lies within 2.5 standard deviations of the model then the current pixel is assumed to be a shadow and is considered as a part of the scene background.

2.1.3 Object Image Segmentation

In the result of the background modelling, a binary mask denoting pixels recognized as belonging to foreground objects in the current frame is obtained. It needs to be refined by the means of morphological processing in order to allow object segmentation [7][10]. This process includes finding connecting components, removing objects being too small, morphological closing and filling holes in regions.

Additionally, an algorithm for shadow removing from the mask using morphological reconstruction is implemented [11]. The morphological reconstruction procedure involves two binary images: a mask and a marker. In the mask image, all pixels belonging to either the moving object or the shadow have value of one, and all the background pixels have zero value. The marker is obtained by applying an aggressive shadow removal procedure to the object detection result, so that all the shadow pixels are removed, some pixels belonging to the moving object may also be removed in this process (the object mask is damaged). The marker is first 'cleaned' by removing isolated pixels, then it is dilated by a structural element which usually has a large size (typically, 9 x 9 structural element is used). The result of marker dilation is then combined with the mask using logical AND operation. As a result, shadows are removed and the moving object masks are properly reconstructed.

2.1.4 Object Detection Experiments

This section demonstrates the performance of the object detection algorithm. During experiments the number of distributions K was set to 5, the α value was 0.0001, and the threshold T was set to 0.5. Video sequences were recorded with 25 fps.

Overall performance of the algorithm is satisfactory. Sample results of object detection and segmentation are presented in Fig. 1. It is seen that employing the shadow detection and removal significantly improves accuracy of binary masks obtained.

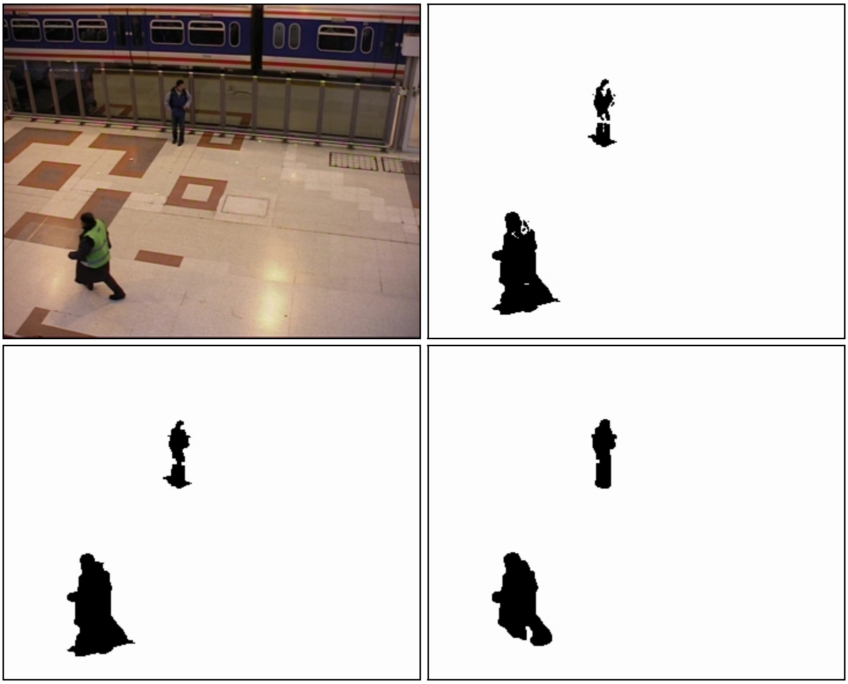


Fig. 1. Sample results of moving object detection and segmentation in the frame number 1678 of the S1-T1-C3 recording from the PETS 2006 [12] set; the first image shows original frames from recorded video sequences; the second image contains the results of background removal without any further processing; the third image shows binary masks after morphological operations; the results in the second and third images were obtained with the shadow detection module disabled, while the fourth image shows the final results of segmentation with moving shadows removed.

2.2 Moving Object Tracking

After a successful moving object detection it is necessary to track all detected objects during their presence in the camera field of view in order to analyse their behaviours, interactions and detect potentially dangerous events. This process involves solving many problems regarding object occlusions.

2.2.1 Applying Kalman Filtering

The Kalman filtering provides a useful approach to tracking objects in video, thus numerous papers discussed this subject [13][14][15]. In the process of tracking, each of the detected moving objects is assigned its own Kalman filter (so-called tracker). The Kalman filter is most of all used to establish proper relation between detected regions that map to moving objects of a current frame, and the real moving objects under analysis [16].

As a result of applying algorithms for moving objects extraction described in previous sections, moving objects represented by rectangles are obtained. Therefore, the state vector of Kalman filters (trackers) may contain the following parameters:

$$x = \begin{bmatrix} x & y & w & h & dx & dy & dw & dh \end{bmatrix}^T \tag{5}$$

where x and y denote the location of the object, w and h are the width and height of the rectangle, dx and dy indicate the change in the object's location during subsequent time intervals, and dw and dh are the changes in the width and height of the rectangle during subsequent time intervals.

The measurement vector adopts the following form:

$$z = \begin{bmatrix} x^b & y^b & w^b & h^b \end{bmatrix}^T \tag{6}$$

which includes: the location, width, and height of the region holding the pixels of the moving object associated with the current tracker.

Thus, transition matrices A and H are as is shown below:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{7}$$

The applied model does not require any control inputs, which results in an input matrix B equal to 0.

The matrices of noise covariance of the process Q and the measurement noise R covariance used in the experiments were constant and equal to:

$$Q = \operatorname{diag}\left(10^{-5}\right) \qquad R = \operatorname{diag}\left(10^{-3}\right) \tag{8}$$

where operator diag(a) denotes the diagonal matrix with values a in its main diagonal.

The process of tracking a moving object has several phases. Each newly detected object is assigned a new tracker with the state vector that is set based on the parameters measured for the region according to the following equation:

$$x_{-1} = \begin{bmatrix} x_{-1}^b & y_{-1}^b & w_{-1}^b & h_{-1}^b & 0 & 0 & 0 & 0 \end{bmatrix}^T \tag{9}$$

In the next frame, the state vector is updated once more, based upon the parameters corresponding to the newly-created object:

$$x_0 = \begin{bmatrix} x_0^b & y_0^b & w_0^b & h_0^b & x_0^b - x_{-1}^b & y_0^b - y_{-1}^b & w_0^b - w_{-1}^b & h_0^b - h_{-1}^b \end{bmatrix}^T \tag{10}$$

The vector x_0 constitutes the initial estimate of the state vector $\hat{x}_0$. The initial estimate of the error covariance P_0 is assumed 1 (one). In the following time intervals, firstly the forward prediction of the state vector of all Kalman filters assigned to the currently existing objects is made. This is done in order to obtain the *a priori* estimate of the location of objects belonging to the current image frame. The next step is to purge trackers whose *a priori* estimate of the state vector contains non-positive or too small values of the object's width and height. Then, it is decided which region of the current image frame is assigned to which one of the tracked objects, which is the mayor part of the algorithm described in the following section. In the final phase, the Kalman filter state vectors of each object are corrected. This is done based on the parameters measurement of the regions holding the pixels of the respective moving objects detected in the current image frame.

2.2.2 Establishing Relations between Moving Objects and Regions

The key action of the tracking algorithm is to associate trackers with the blobs resulting from background subtraction in the current frame properly. A binary matrix M that depicts the relations is created for this purpose,. In this matrix, each tracker-blob pair (where the *a priori* estimate $\hat{x}_k^-$ of the object's state represents the tracker and the measurement vector z_k relates the blob) is assigned 0 or 1, depending on whether the rectangles enclosing the region and the estimated object location have a common part (i.e. whether they overlap or not). As a result a $i \times j$ relations matrix M is created for i trackers and j detected blobs of the current frame. This way of the matrix creation provides a vital simplification of hitherto used procedures.

There are 6 basic types of relations between trackers and regions, each of them requiring some different actions to be taken [16]. These relations and actions are summarized in Tab. 1. The algorithm was tuned to assure the integrity of the traced object in case of faulty background subtraction or when an object temporarily disappears behind an obstacle (e.g. pillar) or to achieve a continuous, valid tracks in situations when objects temporary occlude each other.

Table 1. Possible relations between trackers and regions denoting moving objects and actions taken in each case

Type of relation	Action
Region without a tracker	New tracker created based on the region properties
Tracker without a region	Only prediction phase carried out; if the tracker fails to relate to a region within several subsequent frames, it is deleted
One tracker – one region	Tracker is updated with the parameters of the region
One tracker – many regions	Tracker is updated with the parameters of the rectangle covering all regions
Many trackers – one region	Each tracker is updated with the parameters of the same region.
Many trackers – many regions	Iterative analysis of region-tracker clusters and updating a tracker with the most suitable region based on visual similarity.

In order to relate each tracker to the valid region successfully in case of the many trackers to many regions relation, a two-dimensional colour histogram using a chromatic space of R_cG_c colours was applied for each object. The histogram dimensions were chose as: 16×16 pixels. The values of R_c and G_c in the chromatic space may be derived from an RGB pixel set through the following formulas:

$$R_c = \frac{R}{R+G+B} \qquad G_c = \frac{G}{R+G+B} \tag{11}$$

This means that the values of R_c and G_c represent the share of red and green in the original colour. It is easy to notice that deriving analogously the share of blue is unnecessary, as the 3 shares are altogether equal to one.

The chromatic space defined in that way introduces some losses (it is not possible to revert to the full RGB space of colours). However, the space does not include the information about luminance which causes that an object moving in an unevenly illuminated scene has an identical appearance in the R_cG_c space all the time. For example, for $R_c = 1/3$ and $G_c = 1/3$, it is known that all the compound colours had the same share in the original RGB space, although it is not possible to determine whether it was black, grey, or white. Another advantage of the R_cG_c colour space is the simplicity of its specification and a lesser dimension of histograms based on it.

The degree of similarity between the appearance of the object (the R_cG_c histogram) stored by the tracker, and the appearance (the R_cG_c histogram) of the analysed region is determined through the measurement of correlation. In order to decide that the region and the tracker represent the same object, the correlation must be at least 0,95. The colour histogram is updated only at the time the tracker is created or when it is in the one-to-one relation with a specified region.

2.2.3 Object Tracking Experiments

The experiments show that the developed algorithm works correctly. The performance results are satisfactory, especially when the number of objects (people) who are present in the analysed scene is not large. Sample results are depicted in Fig 2. It is seen, that each tracker (marked by a different colour) follows the right object.

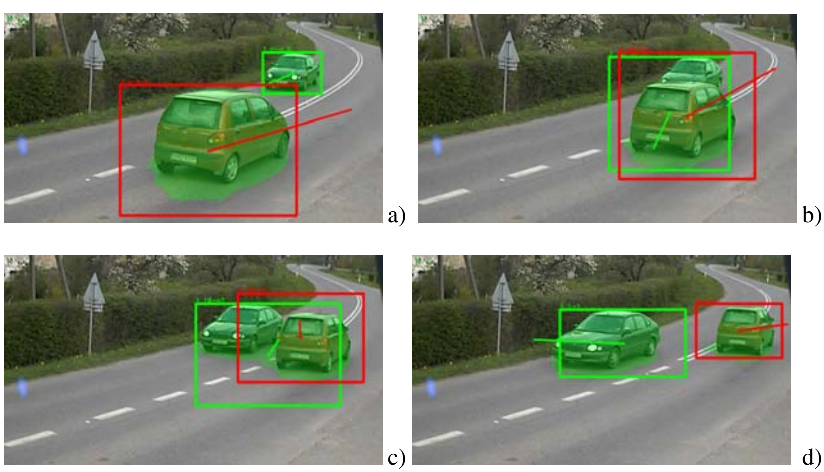


Fig. 2. Frames illustrating continuous tracking of two vehicles passing by each other (with instantaneous motion vectors and green masks denoting results of object detection applied)

Based on the results of moving object detection and tracking, various parameters regarding each object individually and the analyzed scene as the whole, may be easily derived. Such an individual object parameter vector may contain such descriptors as object localization and dimensions, instantaneous and average speeds, various statistical descriptions of object texture etc. Among global scene parameters one may find a number of objects that appeared in and then left camera field of view, their average speed, average ground area covered by moving objects and many more. The set of parameters calculated for each object and for the scene is determined by requirements of higher-level data analyzing modules. Utilization of the parameters for classification of dangerous events provides the scope of Sec 3.1.

3 High Level Image Processing for Detection of Threats

The first stage of processing of camera images in video surveillance systems includes low-level procedures such as background separation, shadow removal and object tracking, providing the basic information about the moving objects in the observed area. At the higher level of image processing, interpretation of events occurring in the camera images takes place. The processing chain may be separated into several modules. First, moving objects detected by the object tracker are examined, parameters describing these objects are extracted and the objects are separated into classes by an object classifier. Next, interactions between each object and other objects are examined. Typically, a set of rules is defined in order to describe typical events occurring in the observed area. An event detector is to test these rules for all the tracked moving objects and if any rule matches, an event is detected.

3.1 Object Classification and Camera Calibration

Before the event detection in camera images is performed, the system has to learn the type of each moving object. In other words, several classes of objects have to be defined and the rules that examine the object parameters and assign the object to one of the classes, have to be created. Typical classes of moving objects include: vehicles, humans, animals and objects. Some subtypes may be defined within each class,. For example: a class of vehicles may be divided into passenger cars, vans, trucks, buses, motorcycles, bicycles, etc.

The object classification module processes data concerning moving objects, provided by the object tracker, and outputs a symbol of the class to which the object was assigned. Two main procedures are performed inside this module. First, the parameters (often called descriptors) of each moving objects are extracted. The main parameters that are used for object classification are related to the size of the object (width, height and proportion of these two measures). Other parameters may describe object's appearance and movement. A vector of parameters used for the classification has to be selected carefully. Increasing number of parameters not always improves the classification accuracy, sometimes it can even deteriorate it. Therefore, a subset of all possible parameters, providing best classification results, has to be selected.

The size of the moving object is almost always used in the classification process. However, it is not possible to classify object based on their size measured in pixels in the camera image, because of the perspective effect. A distant object, with small pixel area, may be in fact larger than the object closer to the camera. Therefore, a method for calculating the real size of the moving object is required. Since the camera object is two-dimensional and the information of the image depth is not available, the size of the object can only be estimated.

3.1.1 Camera Calibration

In order to estimate the real size of the object, a conversion between the camera coordinates (measured in pixels) and the world (physical) coordinates (measured in meters or inches) has to be defined. This can be achieved by means of the camera calibration procedure which is based on measuring the position of characteristic points (calibration marks) both in the camera image and in the real world. One of these points is usually set as an origin of the world's coordinate system. The calibration algorithm processes pairs of the world's coordinates and the image coordinates of each point, in order to calculate conversion coefficients.

Various calibration methods may be found in the literature, but Tsai's method, based on the pin-hole model of perspective projection, is one of the most popular and accurate ones [17]. Given the position of a point in 3D world coordinates, the model predicts the position of the point's image in 2D pixel coordinates [18]. Tsai's model has 11 parameters: 5 internal (intrinsic, interior) and 6 external (extrinsic or exterior) parameters. The internal parameters are related to the camera construction and include a focal length, a lens distortion coefficient, coordinates of the center of radial lens distortion and a scaling factor. The external parameters are related to camera's position and orientation (rotation and translation components for x, y and z axes).

Two modes of camera calibration are available. In the coplanar calibration, it is assumed that all calibration points are situated on the same z plane. If the calibration points differ in height, a more complex non-coplanar calibration has to be used. At least 7 points in the coplanar mode and 11 points in the non-coplanar mode are needed in order to achieve a proper calibration accuracy. An example of the calibration points marked on the camera image and measured in the real world is shown in Fig. 3.

After the calibration process is done, the 3D world coordinates of any point may be converted to 2D image coordinates. The inverse conversion is not directly possible, because the translation from 2D image coordinates to 3D world coordinates is ambiguous. In order to make this conversion possible, a constraint has to be introduced by specifying the height of the point in the world coordinates. However, this height must be known and this is not the case, once only the camera image is available. Therefore, the only method to input the height value to the conversion routine is to apply an estimation procedure.

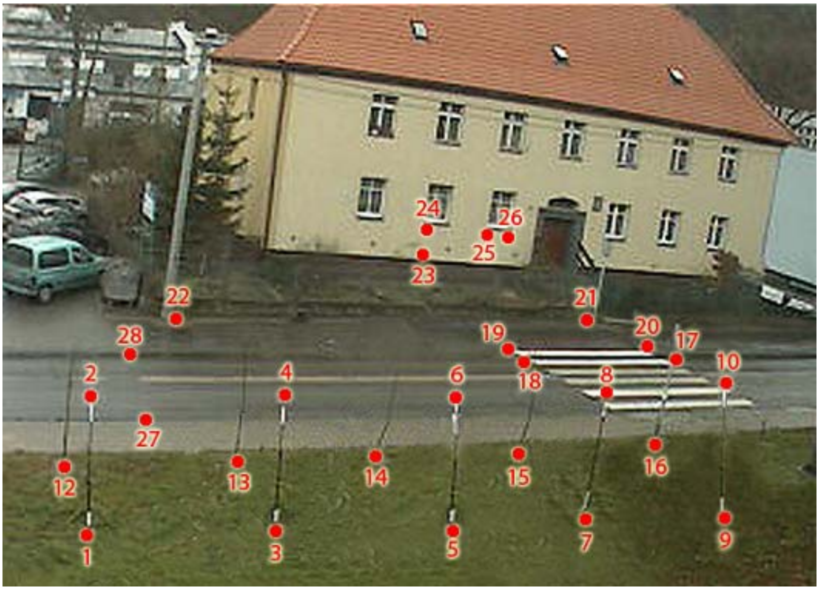


Fig. 3. Example of camera calibration setup. Calibration points (including landmarks and calibration posts) were selected (numbered points in the image) and their positions were measured both in meters (real world measurements) and in pixels (in the camera image).

3.1.2 Estimation of Real Size of the Objects

Each tracked moving object is described by its bounding box (a rectangle). The height and the width of this box may be treated as object's sizes in the image coordinates, provided that this box encloses the whole single object and it does not include other objects, shadows, etc. Therefore, high accuracy of background subtraction and shadow removal algorithms is required. Assuming that the bounding box describes the size of the object in the image coordinates, it is possible to estimate the width and the height of the object using the size of this box. The estimation of the physical width of the object is straightforward, because it may be usually assumed that both the lower corners of the box are situated on the same plane and they have identical z_w coordinate (which may be set to zero in order to simplify the calculations). Therefore, the estimated physical width w of the object is calculated by converting the image coordinates of the lower left and the lower right corners of the box to the world's coordinates and measuring their distance. Estimation of the physical height of the object is much more complicated, because the two points that define object's height (e.g. the lower right and the upper right box corners) are situated on different horizontal planes. One possible method of height estimation is the cast shadow method [19]. First, the object's bounding box is assumed to lie on the ground plane (it is treated as a 'shadow' of the object). The bottom left corner of the box is converted to the world coordinates. Then, a virtual reference point, situated directly h_R meters above the box corner, is selected and converted to image coordinates (Fig. 4). The height h_R is chosen arbitrarily. Next, the cast of the points h_R and h_C (the upper left box corner) on the

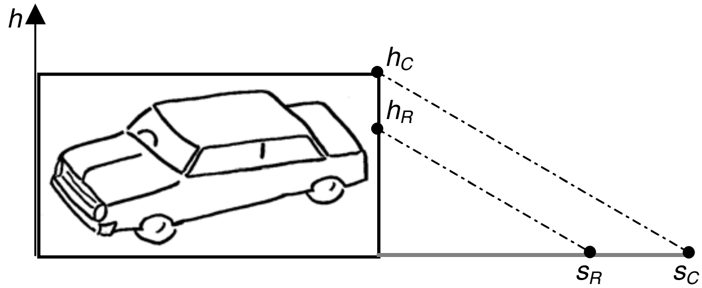


Fig. 4. Illustration of the method of object's physical height estimation

ground are found (noted s_R and s_C, respectively). Knowing the coordinates of three points (h_R, s_R and s_C), it is possible to obtain the height h_C of the object from the proportion.

The results of the size estimation in the consecutive frames may change considerably because of inaccuracies of the background subtraction procedure. In order to smooth the estimation results, some post-processing may be applied, e.g. the output value of the estimated width or height of the object may be calculated as a moving average of the last 10 measurements. In order to avoid errors in size estimation for overlapping objects, a new measurement is added only on the condition that no conflicts occurred in the object tracking procedure,. Additionally, the procedure of size estimation will not provide the correct results unless the measured object frame describes an exact position and a size of the moving object. Therefore, size estimation should only be performed if the object is fully in the image frame, it is not in conflict with any other moving object and the size of the bounding box is larger than the minimum size (this helps to avoid estimation errors for distant, small objects). It should be noted here that the size of the bounding box of the moving object and not the real size of the object is estimated here, which is evident in Fig. 4. Depending on the camera and the moving object orientation, difference between the object's box size and the real object's size may be considerable.

The accuracy of size estimation depends on the number of calibration points, spatial distribution of the calibration point in the camera image and the accuracy of measurements of distances between the calibration points. If the calibration points are concentrated in a part of the observed scene, size estimation will be accurate in the area covered by the calibration marks, but the estimation error will increase as the object's distance from the calibrated area increases. Using a larger number of calibration points, distributed evenly in the whole camera image, provides higher estimation accuracy, on the condition that the distances between the calibration points are measured correctly. Moreover, the position of the calibration points in the camera image also has to be found with high accuracy. This often requires a usage of high resolution camera images. Every inaccuracy in measurement of distance between the calibration points, either in the real world or in camera image, results in inaccurate camera calibration and wrong size estimation. Additionally, for accurate height estimation, non-coplanar calibration has to be used, with calibration points situated at a different height above the ground.

3.1.3 Object Classification

The object classification module, as mentioned previously, operates on a set of rules that examine the object parameters and assigns the object to a class using the defined rules. In order to determine the rules needed for a proper classification, the first a set of test materials (camera video recordings) has to be collected. The moving objects detected in these recordings have to be manually assigned to classes and to the values of selected parameters of these objects need to be calculated. The analysis of these parameters in different classes allows for determining the set of object descriptors and the rules assigning the objects to classes.

The simplest object classifier is based only on the object's size. Such a module may operate on the estimated sizes of the bounding boxes of moving objects, as described in the previous section. The classification is performed using a set of rules, obtained by analyzing the test material. Each moving object in the test recording has to be manually assigned to a class, while its estimated size is measured. By analyzing a set of moving objects parameters, it is possible to assess the range of width and height of the representative objects in each class and to construct the classification rules accordingly. For example, the rule may assign the object to the class 'humans' if the object's estimated width lays between 0.5 m and 2.5 m and its height lays between 1 m and 3 m, another rule may classify the object as 'passenger car' if its width remains between 3 m and 4.5 m and its height remains between 1.5 m and 3 m, etc.

The system assigning objects to classes based only on their size is greatly simplified and may be treated as an illustration of the object classification problem. It will not provide high classification accuracy in complex situations. However, more advanced classification systems, operating on a larger number of rules, classes and parameters, follow the general idea presented here. In order to obtain a proper object classification, additional parameters should be used in the rules. Typical parameters useful for object classification, besides the object's size, are:

- velocity (current, averaged, maximum, acceleration),
- characteristics of movement (changes in direction of movement, frequency of stopping and starting movement),
- size variability (e.g. humans typically exhibit larger changes in 'height to width' proportion than vehicles),
- shape,
- structure (e.g. edge detection results),
- colour,
- texture.

Usually, some post-processing of the classification results (averaging, discarding erroneous values, etc.) is required in order to obtain a set of parameters properly, describing the moving object. It should be also noted that increasing the number of classification parameters may in some cases deteriorate the classification accuracy, if the rule using the added parameter makes an incorrect decision. Therefore, the design of the accurate object classifier requires the thorough analysis of the available parameters, selection of the optimal parameters and definition of rules.

3.2 Event Detection in Video

The main part of the automatic system for analysis of the camera images is the event detector. The task of this module is detection and interpretation of events occurring in the camera images. The event detector works by testing the rules that describe the events, using the parameters obtained from the object tracking and the object classification modules.

The event detection module may be divided into two functional parts, performing the detection and the analysis of the events, respectively [19]. The first part is the *low-level event detector* (LLED), situated lower in the system hierarchy. The LLED uses parameters describing the object's position, physical size, class, statistics, etc. for the detection of low level-events, such as an object entering or leaving the screen, object that stopped or started moving, etc. The task of the *high level event detector* (HLED) is the interpretation of detected low-level events and detection of complex events, such as a car parking in the observed area, a person getting into a building, etc. Both parts of the event detector have their own set of rules. The rules for HLED are defined using terms that are similar to description of events in the natural language, therefore the HLED is positioned as closest to the user interface in the system architecture.

The simple example of the LLED is the module evaluating the position of the object relative to the screen, the class to which the object was assigned to by the object classification module, and the current direction and velocity of the object's movement. Using these descriptors, the following set of rules for detection of low-level events may be formulated:

- The object is entering the screen, crossing its border. The position of the object and direction of its movement are detected. The detected event may be logged using a description like "an object entered the screen at position (x,y) from the left side".
- The object appeared in the area inside the screen (not crossing the border), e.g. car leaving the parked area, a person leaving the door through the building. Example of event description: "a new object appeared in the screen at position (x,y)".
- The object left the screen, crossing its border, e.g. a car driving away from the observed area.
- The object disappeared in the screen, but not leaving the screen margin, e.g. a parked car.
- The object stopped moving, e.g. a car waiting at the traffic lights.
- The object resumed its movement, e.g. a bus leaving the bus stop.
- The object was assigned to a class by the classification module, e.g. "object No. 113 was assigned to a class 'human'.
- The object entered a defined area. For example, an operator may define a 'forbidden zone' in the image. If an object belonging to a selected class enters this area, an event is detected, e.g. "object No. 113 of class 'human' entered zone No. 2".

Detection of the low-level events should not be performed using parameters obtained from the analysis of the current camera image frame only, because short-term variations and errors in image content analysis may result in erroneous detection of events. For example, an error in object tracking may result in an overestimation of the object's width and height for just one camera frame, but it would result in detection of a non-existing event. Therefore, it is important to perform event detection on post-processed (e.g. averaged) parameters describing the moving objects.

Fig. 5 shows an example output of the working LLED module, obtained by the analysis of a series of camera images using a simple set of rules. Since this is a low-level event analysis, no interpretation of the event is performed at this stage, only a 'raw' events are detected and presented in the module's output.

```
13:52:34: Object [25] appeared at (124,112)
13:52:35: Object [25] was assigned to class 'car'
13:52:36: Object [25] started moving
13:52:44: Object [24] of class 'human' left the screen, going left
13:53:06: Object [25] of class 'car' disappeared at (174,128)
13:53:34: Object [26] appeared at (65,112)
13:53:36: Object [26] started moving
13:53:39: Object [26] was assigned to class 'human'
13:54:22: Object [26] of class 'human' left the screen, going right
13:54:40: Object [27] entered the screen from the right
13:54:41: Object [27] was assigned to class 'car'
13:54:51: Object [28] entered the screen from the left
13:54:59: Object [28] was assigned to class 'cars'
13:55:51: Object [27] of class 'car' left the screen, going left
13:56:03: Object [28] of class 'car' left the screen, going right
```

Fig. 5. Example fragment of log from low-level event detection procedure

While the LLED rules operate only on the parameters of tracked objects, the HLED rules also interpret the series of consecutive low-level events in order to detect the high-level events. In other words, the high-level rules describe what the events detected by the LLED mean. The high-level rules may be very simple, for example:

IF	an object of class 'humans' disappeared in the area described as 'the door of the building'
THEN	a person entered the building

In a practical system, the high-level rules are often much more complex, as they interpret a number of detected low-level events occurring in a certain period. An example of the complex rule examining a chain of the consecutive events, is presented below.

IF	object No. 1 entered the screen from the left
AND	object No. 1 was assigned to class 'car'
AND	object No. 1 stopped in the area 'car park'
AND	object No. 1 disappeared at position (x1, y1)
AND	object No. 2 appeared at position (x2, y2)
AND	distance between points (x1,y1) and (x2,y2) is below the threshold
AND	object No. 2 was assigned to a class 'human'
AND	object No. 2 started moving
AND	object No. 2 left the screen, going right
THEN	a car drove from the left side of the screen and parked in the area 'car park', then a person got out of the car and walked away through the right border of the screen

Rules such as the one presented above, may be defined by the operator using an available set of the low-level events or they may be selected from a predefined set. The implementation of the HLED system requires a logical structure (database) that stores the detected low-level events, together with their time signatures and screen positions, and a rule interpretation algorithm that browses the detected events and evaluates all possible rules. The rule interpretation system may operate on a system of simple deterministic rules or it may utilize more complex methods, such as fuzzy logic, that provide much more possibilities for processing of the rules, especially the high-level ones.

Examples of typical high-level events detected in the automatic video content analysis systems, are presented below:

- Intrusion detection – an object of a defined class, entering the 'forbidden zone', e.g. a person trying to cross the motorway, a car in the 'no entry' area, a person climbing the wall, etc.
- Traffic regulations violation – e.g. a car ignoring the red light, a car parking in the restricted zone, a private car in the bus lane, etc.
- Abandoned luggage – an object that is dropped by a person and left without any person in its vicinity for a defined time (important e.g. for airport security).
- Theft – an object disappearing or of its usual position (e.g. a picture removed from the wall in a museum).
- Loitering – a person moving around in the observed area for a defined time.
- Vandalism – detecting the act of destroying objects (e.g. demolishing a bus stop, painting a graffitti).
- People counting – counting a number of persons in the observed area (airports, museums, sport arenas, etc.) and raising an alarm if the number of persons exceeds the threshold.
- Camera sabotage – detection of the cameras that were damaged, switched off, moved, stolen, painted over, etc.

The main problem in event detection is that the system has to cope with the situations in which the event occurs, while the important objects are not visible in the camera image. For example: the abandoned luggage needs to be detected even if the person was obscured by another object while they were dropping the luggage on the floor. Therefore, an advanced processing of detection results gathered from analysis of a large number of frames is required in order to interpret all the important events. Even more complex analysis procedures are needed when the task of the system is to detect the potential events, before they would occur. Such a system needs to analyze the behaviour of the moving objects and compare it to some patterns of typical object's actions in order to predict what will happen next. The discussed systems often implement complex models of behaviour of humans and other moving objects. Specifically, modelling human behaviour, together with pose estimation algorithms, helps to detect and interpret complex events such as fights, assaults, thefts, riots, etc. For example: a person loitering around the parked bicycle may be a potential thief, therefore the system operator should be notified so that they may investigate the situation.

It is evident that the system for automatic event detection can never reach a 100% accuracy in practical applications. Therefore, the system has to be carefully tuned up in order to obtain a satisfactory balance between false positive results (detecting an event that did not occur) and false negative ones (missing an important event). In practical applications, a number of false negative results should be kept as low as possible, so that no important events are missed by the system. A certain level of false positive results is acceptable, as long as it does not make the system operator to ignore the system alerts, because they would think that the system is overreacting.

3.3 Object Tracking with PTZ Cameras

In real video surveillance systems, multiple cameras are used for monitoring of different areas that may be separated or may overlap. Video monitoring systems use stationary wide-angle cameras (with constant field of view) and pan-tilt-zoom (PTZ) cameras (where field of view may be moved and zoomed). The main problem here is that the results of video analysis obtained from multiple cameras have to be combined in order to achieve an accurate event detection. For example, if any object leaves a field of view of one camera and, after some time, it enters the field of view of another camera, the system has to recognize it as the same object.

The problem of object tracking using multiple cameras is still a topic of research and development, with various approaches to this task [20][21][22][23]. In the automatic multi-camera monitoring system there are two main problems that need to be solved. First, a proper configuration of the cameras has to be chosen. Wide-angle cameras may be used for efficient object detection and tracking but they do not have zoom capabilities. On the other hand, PTZ cameras provide a detailed full-screen view of the object, however a detection and tracking of other moving objects is not possible at the same time. Therefore, a solution to this problem can be a introduced to the camera system. The wide-angle cameras are used for object detection and tracking, as well as for event detection. In case an event is detected, PTZ cameras are used to provide a detailed view of the selected moving object and to track its movement. Therefore, a second problem to be solved is related to positioning and zooming the view of PTZ cameras in order to track a particular object. Moreover, if the object leaves the field of view of a camera, another camera has to capture this object and should continue its tracking.

This section presents a novel approach to the described problem, utilizing modern technologies [24]. The tracked object is equipped with a GPS receiver and a mobile terminal running a client service that continuously transmits the position of the object to the server. The task of the server is to convert geographical coordinates of the object to parameters of PTZ cameras, to select the cameras that are able to track the object at the moment and to adjust the field of view of these cameras appropriately.

Efficient object tracking with PTZ cameras requires two steps to be performed. First of all, an algorithm for converting of object's GPS position to appropriate setting of a PTZ camera (including prediction of the object's movement to compensate for the delay related to signal transmission and processing) needs to be implemented. Furthermore, each camera must be calibrated in order to be able to lock on a moving object with enough high accuracy. And finally, a method for tracking the same object with multiple cameras, considering their limited areas of coverage, needs to be developed.

3.3.1 Conversion of Object's GPS Position to PTZ Camera Settings

In order to fix a PTZ camera on a nearby object based on its known GPS position it is necessary to calculate pan, tilt and zoom parameters for a dome camera in the way that guarantees that an object of known longitude and latitude will be present in a video stream. Because of a very low accuracy of altitude data from a GPS receiver, it is assumed that all objects move on the same altitude equal to the ground level.

The algorithm has 4 parameters related to the camera: its longitude lon_C, latitude lat_C (both in degrees), height above the ground level h_C and pan offset p_{off}, which is defined as a clockwise distance (in degrees) between the north and a the camera zero pan position. These parameters are derived during the calibration process. Input data for the algorithm include 4 parameters measured by a GPS receiver moving with an object being tracked: longitude lon and latitude lat of the object (in degrees), its speed v and bearing br (in degrees).

In its first stage, the algorithm GPS coordinates of an object are translated to the local, Cartesian 3D coordinate system [24]. In this system camera position is denoted as $(0, 0, h_C)$ and object position is described as $(x, y, 0)$.

There is a significant delay in the system of tracking moving objects with PTZ cameras, caused by the GPS receiver, data transmission (e.g. with GPRS protocol), computing time and executing PTZ command by the camera. This delay must be compensated in order to assure that a fast-moving object is always present in a video frame (preferably: in its center). The delay compensation is performed by setting the PTZ camera to the predicted position $(\hat{x}, \hat{y}, 0)$ of the object. The prediction time should be equal to the total delay in the system. In the current implementation, a linear predictor is used that estimates object position based on its speed v and bearing br reported by a GPS device. Based on the camera location and the target location, the pan and tilt parameters for the PTZ camera may be obtained with standard trigonometric operations [24].

The last camera parameter setting, the zoom, is set based on the object's distance. The closer the object is to the camera, the smaller is the zoom value. This approach assures that object dimensions in video remain more or less constant.

3.3.2 System Calibration

In order to achieve satisfactory accuracy of moving object tracking, 4 camera parameters (lon_C, lat_C, h_C, p_{off}) must be defined with a very high precision. Our initial experiments proved that simple measurement of camera position with a GPS receiver does not provide the required precision. Also, determining camera's altitude above the ground might be problematic. Theoretically, the easiest parameter to obtain is the camera pan offset, however the precision required (especially when tracking distant objects) makes any direct measurement very difficult.

Therefore, a one-time optimization approach was chosen to estimate camera parameters with a sufficient precision. Initial approximations of four parameters are obtained with direct measurements. Then, all parameters are further tuned up during a two-stage, non-linear optimization process, which iteratively minimizes cost functions representing the difference between pan and tilt values calculated by the algorithm and the ones measured directly [24]. Localizations of N points scattered equally in a camera vicinity were measured with a GPS receiver for this purpose. In the same time, pan and tilt values were obtained from a camera pointing at each spot. The camera zoom does not influence pointing accuracy, thus this parameter is omitted in the optimization process.

The number of calibration points should not be too small, because reference data gathered with GPS receiver at each point are often inaccurate and their precision is unknown., 10 calibration points for each camera were used during our experiments.

3.3.3 Multi-camera Tracking

It is possible to track an object by many cameras at the same time, provided that the object may be “seen” by a camera. Area of coverage (AOC) for each camera is defined by an operator. This area denotes regions in the camera vicinity that can be observed by a camera and are not obscured, e.g. by buildings. AOC for each camera is defined as a polygon (or set of polygons) with vertices described as geographical coordinates. This facilitates creating AOC for each camera and allows an operator to use any common mapping software.

Whenever an object being tracked enters AOC of any camera i.e. its GPS position is contained inside any of polygons forming AOC of the camera, the camera is automatically pointed at the moving object and tracks the object as long, as it stays in its AOC.

4 Summary

Video processing algorithms were discussed and some experimental results presented in the chapter. The work described here aims at establishing a new quality of urban agglomeration monitoring. Next step in the planned work is to develop a framework comprising aforementioned processing methods, and also geo-positioning information processing for event display on the city map, and a communication software for dispatching the alarms to security forces.

Acknowledgments. The research leading to these results has received funding from the European Community‘s Seventh Framework Program (FP7/2007-2013) under the grant agreement N° 218086.

Resources

A Commercial systems (hardware + software)

VideoSphere, *www.marchnetworks.com*
Opax, *www.opax.com*
Aimetis Symphony, *www.aimetis.com*
MATE Intelligent Video, *www.mate.co.il*
IntelliVision, *www.intelli-vision.com*
XProtect Analytics, *www.milestonesys.com*
FACE-VACS, *www.cognitec-systems.de*

B Key Books

Gonzalez RC, Woods RE (2008) *Digital Image Processing (3rd Edition)*. Prentice Hall

Bigun J (2006) *Vision with Direction. A Systematic Introduction to Image Processing and Computer Vision*. Springer

Nillson F (2008) *Intelligent Network Video*. CRC

Davies ER (2004) *Machine Vision : Theory, Algorithms, Practicalities*. Morgan Kaufmann

C Key Journals

IEEE Transactions on Image Processing, IEEE
Signal Processing: Image Communication, Elsevier
Image and Vision Computing, Elsevier

D Key International Conferences/Workshops

PETS - IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, *www.cvg.rdg.ac.uk/PETS2009/*

North American Video Content Analysis Conference, *www.imsconferences.com/vcausa/*

Video Content Analysis Conference – Europe, *www.imsconferences.com/vcaeurope/*

E Software

OpenCV, Open Computer Vision Library, free, open source, C++ libraries for video content analysis, *opencv.willowgarage.com*

Axis Camera Station, a comnprehensive video management software for monitoring, recording, playback and event management, *www.axis.com*

ZoneMinder, a free video camera security application suite (Linux), *www.zoneminder.com*

F Web Pages

Keith Price's Annotated Computer Vision Bibliography: http://iris.usc.edu/Vision-Notes/bibliography/contents.html

References

[1] Li, H., Ngan, K.: Automatic Video Segmentation and Tracking for Content-Based Applications. IEEE Communication Magazine 45(1), 27–33 (2007)

[2] Liu, Y., Zheng, Y.: Video Object Segmentation and Tracking Using y-Learning Classification. IEEE Trans. Circuits and Syst. For Video Tech. 15(7), 885–899 (2005)

[3] Konrad, J.: Videopsy: Dissecting Visual Data in Space Time. IEEE Communication Magazine 45(1), 34–42 (2007)

[4] Yang, T., Li, S., Pan, Q., Li, J.: Real-Time and Accurate Segmentation of Moving Objects in Dynamic Scene. In: ACM Multimedia-2nd International Workshop on Video Surveillance and Sensor Networks, New York, pp. 10–16 (2004)

[5] Stauffer, C., Grimson, W.: Learning patterns of activity using real-time tracking. IEEE Trans. on Pattern Analysis and Machine Intell. 22(8), 747–757 (2000)

[6] Elgammal, A., Harwood, D., Davis, L.: Non Parametric Model for Background Subtraction. In: ICCV Frame-rate Workshop (1999)
[7] Dalka, P.: Detection and Segmentation of Moving Vehicles and Trains Using Gaussian Mixtures, Shadow Detection and Morphological Processing. Machine Graphics and Vision 15(3/4), 339–348 (2006)
[8] Czyzewski, A., Dalka, P.: Visual Traffic Noise Monitoring in Urban Areas. International Journal of Multimedia and Ubiquitous Engineering 2(2), 91–101 (2007)
[9] Horprasert, T., Harwood, D., Davis, L.: A statistical approach for real-time robust background subtraction and shadow detection. In: Proc. of IEEE Frame Rate Workshop, Kerkyra, Greece, pp. 1–19 (1999)
[10] Dougherty, E., Lotufo, R.: Hands-on Morphological Image Processing. SPIE Press, San Jose (2003)
[11] Xiu, L., Landabasso, J., Pardas, M.: Shadow removal with blob-based morphological reconstruction for error correction. In: Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. II–729–II–732 (2005)
[12] PETS 2006 – a collection of test recordings from the Ninth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, New York, USA (2006)
[13] Funk, N.: A Study of the Kalman Filter applied to Visual Tracking University of Alberta. Project for CMPUT 652 (2003)
[14] Martínez-del-Rincón, J., Herrero-Jaraba, J.E., Gómez, J.R., Orrite-Uruñuela, C.: Automatic left luggage detection and tracking using multi-camera UKF. In: Proc. 9th IEEE Internat. Workshop on Performance Evaluation in Tracking and Surveillance (PETS 2006), New York, USA, pp. 59–66 (2006)
[15] Lv, F., Song, X., Wu, B., Kumar, V., Nevatia, S.: Left-Luggage Detection using Bayesian Inference. In: Proc. of 9th IEEE Int. Wrokshop on Performance Evaluation of Tracking and Surveillance, New York, USA, pp. 83–90 (2006)
[16] Czyzewski, A., Dalka, P.: Examining Kalman filters applied to tracking objects in motion. In: Proc. of 9th International Workshop on Image Analysis for Multimedia Interactive Services, Klagenfurt, Austria, pp. 175–178 (2008)
[17] Tsai, R.: A Versatile Camera Calibration Technique For High Accuracy 3d Machine Vision Metrology Using Off-The-Shelf TV Cameras And Lenses. IEEE J. Robotics Automat. RA-3(4), 323–344 (1987)
[18] Willson, R.: Tsai Camera Calibration Software, `http://www.cs.cmu.edu/~rgw/TsaiDesc.html`
[19] Szwoch, G., Dalka, P., Czyzewski, A.: Objects Classification Based on Their Physical Sizes for Detection of Events in Camera Images. In: NTAV/SPA, Poznań, pp. 15–20 (2008)
[20] Cai, Q., Aggarwal, J.K.: Tracking human motion using multiple cameras. In: Proc. 13th Int. Conf. on Pattern Recognition, vol. 3, pp. 25–29 (1996)
[21] Lee, L., Romano, R., Stein, G.: Monitoring activities from multiple video streams: establishing a common coordinate frame. Proc. IEEE Trans. Patt. Anal. and Mach. Intell. 22, 758–767 (2000)
[22] Calderara, S., Vezzani, R., Prati, A., Cucchiara, R.: Entry edge of field of view for multicamera tracking in distributed video surveillance. In: Proc. IEEE Conf. Advanced Video and Signal Based Surveillance, pp. 93–98 (2005)
[23] Zhou, Q., Aggarwal, J.K.: Object tracking in an outdoor environment using fusion of features and cameras. Image and Vision Computing 24, 1244–1255 (2006)
[24] Dalka, P., Ciarkowski, A., Szczuko, P., Szwoch, G., Czyżewski, A.: Surveillance Camera Tracking of Geo positioned Objects. Paper submitted for the 2nd International Symposium on Intelligent Interactive Multimedia Systems and Services, Mogliano Veneto, Italy (2009)

Chapter 12

Enterprise Attention Management

Stelios Karapiperis[1], Nenad Stojanovic[2], Darko Anicic[2], Dimitris Apostolou[1], and Dimitris Despotis[1]

[1] University of Piraeus, 80 Karaoli & Dimitriou St., 185 34 Piraeus, Greece
{skarapas,dapost,gxen,despotis}@unipi.gr
[2] FZI at the University of Karlsruhe, Haid-und-Neu Strasse 10-14, 76131 Karlsruhe, Germany
nstojano@fzi.de

Abstract. In this chapter we give an overview of technologies and tools that can support the management of attention of employees. Moreover, we present a novel approach for managing users' attention in knowledge intensive organisations which goes beyond informing a user about changes in relevant information towards proactively supporting the user to react on changes. The approach is based on an expressive attention model, which is realized by combining ECA rules with ontologies. We present the system's architecture, describe its main components and present early evaluation results.

Keywords: Attention management, Semantic technologies.

1 Introduction

Success factors in knowledge-intensive and highly dynamic business environments include the ability to rapidly adapt to complex and changing situations and the capacity to deal with large quantities of information of all kinds. For knowledge workers, these new conditions have translated in acceleration of working performance, multiplication of projects in which they are involved and increased collaboration with colleagues, clients and partners. Knowledge workers are overloaded with potentially useful and continuously changing information originating from a multitude of sources and tools. A significant part of a knowledge worker's day can be occupied with searching and looking for information. Moreover, they have to interact with various people from different organisations and in different contexts. In addition, software tools such as email and instant messaging have added to the whole complexity and have created new interruption channels for knowledge workers.

In order to cope with changes in the business environment, the attention of knowledge workers must be always paid on the most relevant information sources. Indeed, a basic need of knowledge workers is to be up to date with information while facing an information overload situation. In other words, the issue is how to select those information resources whose reading will give most benefits to the reader. Moreover, agility and proactivity can be useful in an environment of knowledge workers with overburdened memories: The computer should know what a knowledge worker works

G.A. Tsihrintzis and L.C. Jain (Eds.): Multimedia Services in Intel. Environ., SIST 3, pp. 263–282.
springerlink.com

with, anticipate her or his information needs and show him/her relevant information resources as soon, or even before they need them.

According to Davenport and Beck [7], attention is defined as a "focused mental engagement on a particular item of information". Moreover, attention refers to the process of selecting and selectively processing of information, which is required because the brain has a limited information processing capacity [4]. In an enterprise context, attention management refers to supporting knowledge workers focus their cognitive ability on a particular organisational task and on the information resources required to accomplish it. In particular support is required for searching, finding, selecting and presenting the most relevant and up-to-date information without distracting workers from their activities. Information retrieval systems have provided means for delivering relevant information to users. The main issue with existing systems is that they do not cope explicitly with the information overload, i.e. it might happen that a knowledge worker "overlooks" important information.

The goal of attention management systems is to avoid information overload and to provide notifications about new and changed, relevant information [7]. Moreover, the frequently changing environment requires not only very effective systems for alerting knowledge workers that some relevant piece of information has appeared or has been changed, but also effective recommendation how to deal with these changes.

Our approach for managing attention in a corporate environment is not just to support receiving new relevant information proactively, but also to enable relevant reaction on this information (i.e. on a change in general). Such an action can be an already predefined workflow, but also ad-hock generated procedures according to the currently available knowledge/experience. In that sense, we outline an Enterprise Attention Management System (EAMS) which goes beyond informing proactively a user that relevant information has been changed, toward proactive preparing and supporting the user to react on that change. Our approach puts forward a comprehensive reasoning framework that can trigger a knowledge base in order to find out the best way to react on a change. We base such a framework on a combination of ECA (event-condition-action) rules and ontologies.

The chapter is organized as follows: in the second section we analyze motivating examples, we derive requirements and we outline a framework for an enterprise attention management system, in the third section we present the architecture of a prototype attention management system encompassing various functionalities to address relevant attention-related issues and in the fourth section we present existing technologies and tools aiming at supporting management of attention. We conclude by suggesting outlets for future research and development in information technology for the purpose of managing users' attention in knowledge-intensive environments.

2 Motivating Examples, Requirements and Framework

In this section we outline motivating scenarios and we elaborate on the requirements that stem from the scenarios. Moreover, we present our framework for enterprise attention management.

2.1 Example Scenario 1

In order to keep local laws aligned with the federal laws, the mayor in a municipality has to react on any changes and updates in the federal laws. In this process a General

Binding Regulation (GBR) is being developed by legal experts in the local administration, consulted by a group of experts. Together they discuss about ways to implement the federal law in the local administration. The outcome of this discussion is a draft GBR, which thereafter is open for public deliberation. Citizens can provide comments on the draft GBR and the head of legal department together with the experts assess the comments received. Next, the revised GBR is submitted to the local councillors for approval. If the GBR is approved, it is then signed by the mayor and published.

Since changes in the federal law are announced in the federal governmental web portal, a system for alerting the mayor about new laws is needed. However, the mayor might not be interested in all possible changes, but rather only in some that are relevant for her/his municipality. Moreover, it is not straightforward how the mayor should start the deliberation process upon detection of a new law (i.e. what additional information should be defined, like what is the timeline for the public deliberation). In order to be supported in this decision making process, the mayor would need information related to the specific context, like the information about previous GBRs from this or a related domain. Moreover, relevant information might be hidden or difficult to extract, like that there are not many experts who are familiar with the selected domain in this municipality. This information is needed to enable the mayor to react on the received alert properly. However, the goal would be to provide this information automatically.

2.2 Example Scenario 2

An engineer and his team colleagues are working in the engineering department of a mid and far distances international and transatlantic airline. Their duties are carrying out non-routine maintenance issues arising from airplanes' flights. At any time a non-routine malfunction can trigger an alert into his team concerning a mechanical, electronic or electric issue or any combination of them.

Because there are no approved procedures in standard aircraft maintenance program for non-routine repairs, the engineer's role is to investigate and repair the malfunction in a timely and safely manner. For the airline, every hour on the ground means lost revenues and every repair on the aircraft means consulting and approval by the manufacturer. To improve his team efficiency in performing any reliable non-routine maintenance, he must consider thoroughly all available information sources. Mandatory sources are:

- The airplane's official documentation with up-to-date maintenance procedures.
- Manufacturer's service engineering database with up-to-date service bulletins, service letters, airworthiness directives concerning this particular issue.
- The airline's database with all previous historical data relative to similar cases.
- Other alliance aircraft owners' databases with relative content.

Unofficial but valuable sources are the airline's portal with everyday expertise, the community of practice of aircraft technicians with the best troubleshooting practices and other databases from minor independent aircraft repair stations. By leveraging all possible pieces of related information, the engineering team tries to build a robust pool of knowledge that will ensure elimination of unexpected delays while retaining safety above all considerations.

2.3 Example Scenario 3

An ambitious journalist of a regional electronic newspaper "ICU Daily" is responsible for issuing the politics columns. Her aim is to cover local politics news along with the most interesting national news as well as news from the international scene.

The journalist and other staff of "ICU Daily" gather and generate a daily record of events. This process is time consuming and requires attention and capacity to locate, organize and prioritize the most critical and interesting news. Information is the lifeblood of the "ICU Daily" and as such the journalist needs access to news feeds and up-to-date content concerning the local community. Her prime concern is to ensure that news arriving from any kind of digital media such as web sites, blogs, her associates' e-mails and the like are located as soon as they are released. Moreover, depending on the type of news she receives, she has to use the appropriate template to write her commentary and to send it to the relevant column editor.

2.4 Requirements and Framework

This subsection summarizes the requirements for an Enterprise Attention Management system that can support the aforementioned scenarios. We define three basic requirements for such a system:

1. Flexible modelling of information in order to enable focusing of attention on different abstraction levels. For example, a user interested in information about pets should be alerted for new information about domestic animals (scenario 1). Another issue is modelling the usage of information by a community of users in order to stimulate sharing of implicit knowledge. For example, technicians looking for front tyre pressure of an airplane must be proactively fed with the rear wheels' pressure, as most technicians are interested in both in most maintenance situations (scenario 2).
2. Context-awareness in order to support a particular decision making process. For example, new law about animals triggers different alerts in the GBR and in another business process (scenario 1). Another example is that a new and approved service bulletin should be immediately sent from the airplane manufacturer to all known customers. Moreover, a certain part once rejected from the maintenance list of a service unit should be rejected from all databases (scenario 2).
3. Management of preferences for enabling efficient extraction of interesting information. In particular, there is a need for an expressive formalism for the description of preferences, including when to alert a user, but also how to react on an alert. For example the journalist (scenario 3) ranks highly important news about activities of the mayor and anticipates these headline news to be printed in bold typeface. Another example from scenario 2 is that the engineering team should be immediately alerted with a warning signal once a manufacturer identifies a hazardous device or part.

Figure 1 presents the general framework for Enterprise Attention Management that builds on the aforementioned requirements.

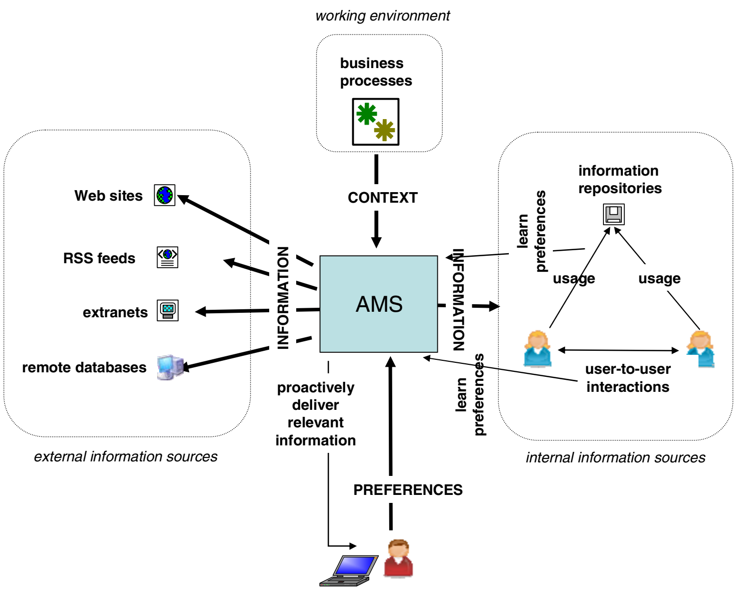


Fig. 1. Enterprise Attention Management Framework

Our framework is developed along 3 axes:

1. Information represents all relevant chunks of information that can be found in the available information repositories and sources. In the business environment of an organisation, sources of information can be both internal and external to the organisation. Moreover, information can be represented either formally (e.g. using information structuring languages such as XML) or informally. Finally, information may be stored in structured repositories such as databases that can be queried using formal languages or in unstructured repositories such as discussion forums.
2. Context defines the relevance of information for a knowledge worker. Detection of context is related to the detection of the user's attentional state that involves collecting information about users' current focus of attention, their current goals, and some relevant aspects of users' current environment. The mechanisms for detection of user attention that have been most often employed are based on the observation of sensory cues of users' current activity and of the environment; however others, non-sensory based, mechanisms also need to be employed to form a complete picture of the user's attentional state [13].
3. Preferences enable filtering of relevant information according to its importance/relevance to the given user's context. In other words, the changeability of resources is proactively broadcasted to the users who can be interested in them, in order to keep them up to date with new information. Users may have different preferences about both the means they want to be notified and also about the relevance about certain types of information in certain contexts. User preferences can be defined with formal rules or more informally by means e.g., of

adding keywords to user profiles. Moreover, even when employing mechanisms capable of formalizing the users' preferences, a certain level of uncertainty about users' preferences will always remain. For this reason, dealing with uncertainty is an important aspect of attention management systems. Equally important is the way preferences can be derived: by explicitly specifying them or by learning techniques.

3 A Prototype Enterprise Attention Management System

The objective of this section is to describe a prototype EAMS that aims to support the requirements outlined above. The overall objective of the proposed EAMS is to support knowledge workers always keep their attention focused on their current tasks by pre-selecting and feeding them with the most relevant information resources to complete their tasks in a way that will not disturb them more than they prefer and by to support them undertake actions relevant to the information received.

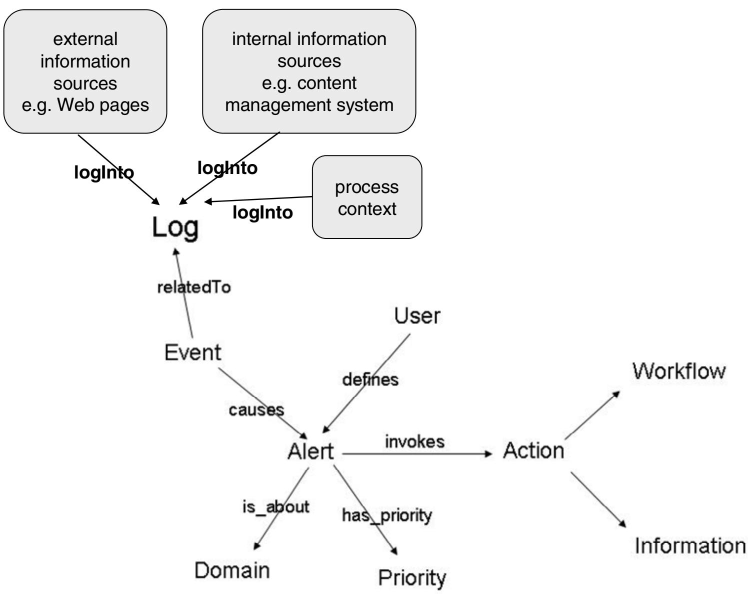


Fig. 2. Conceptual Attention Model

Figure 2 presents the conceptual model behind the proposed EAMS. The model assumes that the interactions between users and internal information sources are logged including the business context (e.g., business process) in which the interactions happened. Some log entries can be defined as Events that cause Alerts, which are related to a user, a problem domain and associated to a priority level. Every Alert invokes Actions, that can be purely informative (i.e. an information push) or executable (i.e. to execute a business process).

In the core of our approach are ECA (Event – Condition – Action) rules; their general form is:

ON *event* AND *additional knowledge*, IF *condition*
THEN DO *something*.

This type of rule is understood as follows: whenever an event occurs, the rule interpreter performs the actions of all matching rules whose conditions hold.

Relevant events are triggered by interactions taking place in enterprise information systems, such as a Content Management System (CMS) or by event sensors that e.g., detect changes in Web pages. Ontologies are used to model and store change related information: The log ontology models change events, following a generic and modular design approach, while the Information ontology contains the domain concepts and relations about which we want to express preferences, such as documents and forum messages. In the following, we describe the log and information ontologies as well as the main components of the proposed Enterprise Attention Management System in detail.

Information Ontology. The Information ontology (an indicative example is shown in Fig. 3) contains the concepts and relations about information resources for which we want to express preferences, such as documents and forum messages. On the top level we have separated the domain concepts from value types.

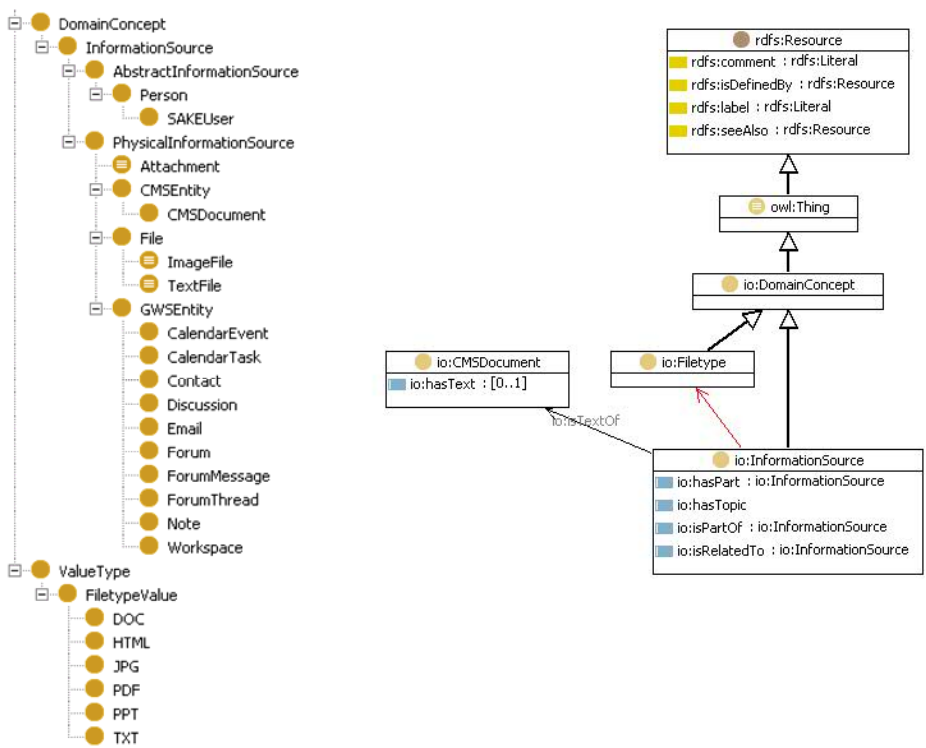


Fig. 3. Information Ontology, Class hierarchy (left), Class diagram (right)

In the InformationSource sub-tree we differentiate between information sources which are of an abstract nature (such as persons), external information sources such as Web pages and RSS feeds, and information sources which physically exist in the organisation, such as documents, forums or e-mails. This FiletypeValue class represents the file type (indicated by the file extension) of a document, for example PPT, PDF, DOC, etc. Note that one subclass of filetype can describe multiple file extensions, such as JPG can be a .jpg or .jpeg file.

Log Ontology. There are many sources of changes that can affect an information resource, like adding, removing, deleting a new document or starting a new discussion. The Log ontology is used for representing these changes in a suitable format. There are four subclasses of Event: AddEvent, RemoveEvent, ChangeEvent and AccessEvent.

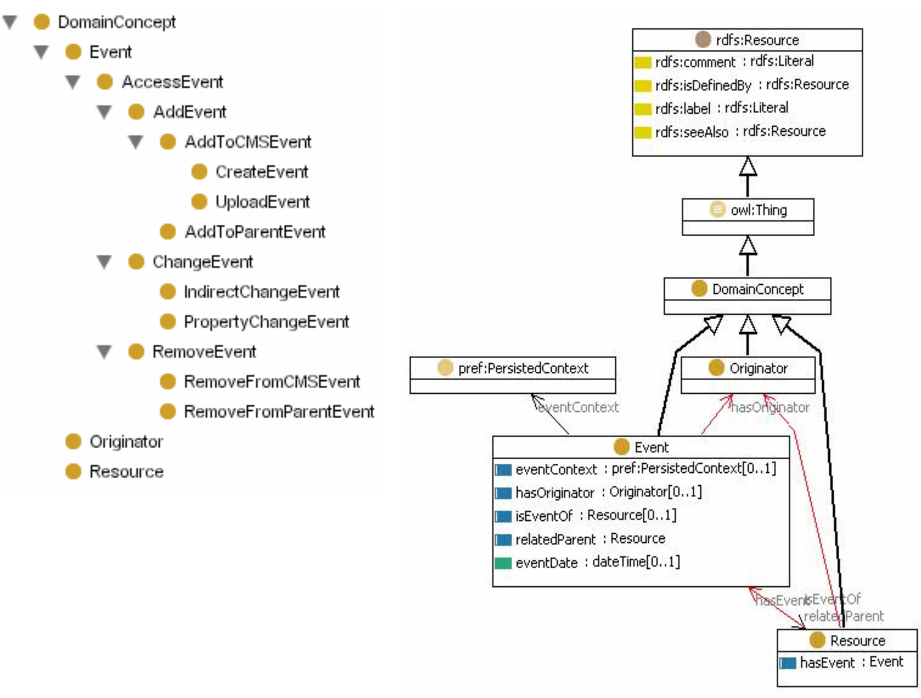


Fig. 4. Log Ontology, Class hierarchy (left), Class diagram (right)

AddEvent is responsible for the creation of new events, e.g. a new document has been added to a CMS (Fig. 4). RemoveEvent is dedicated to the deletion of the existing elements from the system, like the deletion of a document from a CMS. ChangeEvent is responsible for the modification of an existing individual, e.g., the change in the name of the author of a document. It consists of: PropertyChangeEvent, meaning that some properties of an individual have changed and IndirectChangeEvent, meaning a change caused by some other event.

AccessEvent is dedicated to the access of an existing individual. It represents a very broad class of events like reading a document, for which is very complicated to define the semantics clearly. For example, did someone who opened the document and closed it after five minutes, read the document or just opened, considered it as not interesting, but forgot to close it immediately?

We differentiate subclasses AddEvent and RemoveEvent by addition/removal of resources to/from a CMS and by addition/removal of a resource to/from a parent/child relationship using the isPartOf property. AddToCMSEvent is further differentiated by either creating a resource within the system or uploading an existing resource. For ChangeEvents, we distinguish between changes of the resource's properties (e.g. metadata) and changes which are caused by some other event.

Properties of an event are the resources the event relates to, the user who originated the event, a timestamp when the event occurred, an optional description of the event and a copy of the current runtime context. In the case of ChangeEvents we add the names of the resource's changed properties, and optionally link to another event which caused this ChangeEvent.

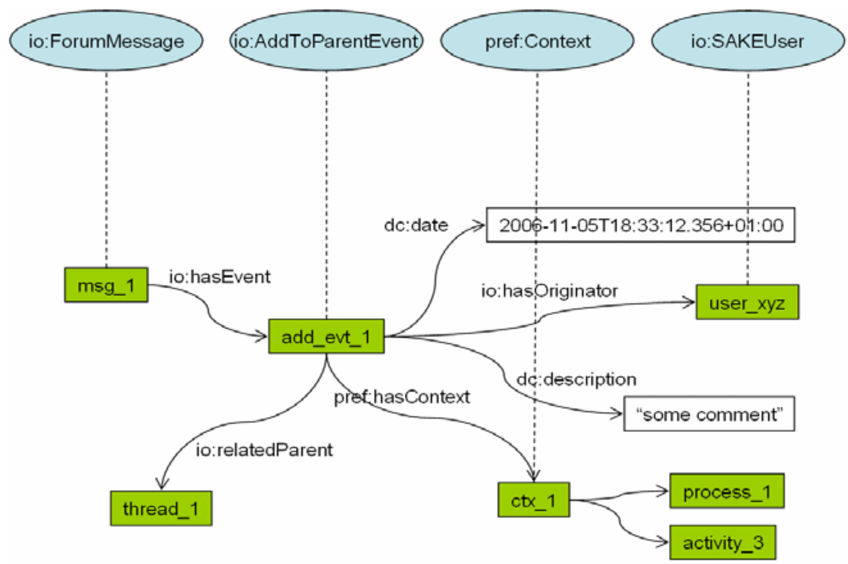


Fig. 5. AddToParentEvent for adding a new message to a thread

The following considerations are assumed:

- In order to see which aspect of the resource has changed, we associate the names of the changed properties with the event.
- If the event has been caused by another event (e.g. the modification of a thread has been caused by the addition of a new message, see Fig. 5), then an IndirectChangeEvent is generated which links to the original event via the causedBy property.

Special attention has to be paid if we define events for "compound" resources, i.e. resources which have child resources. Consider a forum in a groupware system for instance: There we have forums, which consist of multiple threads which consist of multiple messages. Now, imagine that someone adds a new message to a thread. It is clear that by adding a new message to a thread we can consider the thread as changed, thus creating an AddToParentEvent for the message and an IndirectChangeEvent for the thread.

In order to resolve this issue, we define the following as the default behaviour:

- Compound resources state their parent/child relationship by the property isPartOf or a sub-property thereof.
- For compound resources, a ChangeEvent will be generated if (i) the properties of the resource itself change (i.e. title of a forum thread changes) ==> PropertyChangeEvent; or if a child object has been added or removed (e.g. adding a new message to an existing thread) ==> IndirectChangeEvent.
- The modification of a child object does not result in a modification (i.e. IndirectChangeEvent) of the parent.
- The developers programmatically create only the most basic event, e.g. a PropertyChangeEvent or a AddToParentEvent. SWRL rules decide whether this event triggers an IndirectChangeEvent or not as described in the following paragraph. IndirectChangeEvents are never created by the developers.

We do not hard-code the propagation of events from child to parent, instead we define them in SWRL rules, such as:

$$\begin{array}{l} addToParentEvent(?E) \wedge resource(?RES) \wedge hasEvent(?RES, ?E) \wedge \\ relatedParent(?RES, ?RES2) \wedge swrlx: createIndividual(?E2) \Rightarrow \\ indirectChangeEvent(?E2) \wedge hasEvent(?RES2, ?E2) \end{array}$$

Default rules state that the addition/removal of a child object triggers a ChangeEvent for the parent object. However, in order to be more flexible, we could also state that the modification of a specific child object also causes the modification of its parent. Note that in this way, we may use events to specify more complex events (e.g., *indirectChangeEvent*). Those complex events are created using SWRL[1] (Semantic Web Rule Language) rules and a number of built-in predicates supported by KAON2[2]. Although realised in a declarative way, Complex Event Processing (CEP) in our prototype is still limited, and it is subject of our future work. Particularly, we will continue developing declarative CEP. The advantage of such an approach is that definition of a complex event may easily be altered by changing only a logical rule.Further on, inconsistencies in CEP are handled by means of logic.

The business context is derived using the Context Observer, a component that links to enterprise systems such as workflows and extracts the current business process,

[1] http:// www.w3.org/Submission/SWRL/

[2] http://kaon2.semanticweb.org/

activity and task the user is working on. The business context describes the situation which a user is currently present in, and is utilized for derivation of information resources based on context-sensitive preferences.

The EAMS proactively delivers relevant information resources to users and supports them undertake actions relevant to the information received. These capabilities are realised by matching the business context on one side, and user's preference rules on the other side (preference rules are described below). Relationship between the business context and user defined preferences is handled via the *validIn* relation in the preference ontology, Figure 6 (e.g., particular preference rule is *validIn* a certain context). The preference ontology is typically imported by another ontology which maps its own concepts to this ontology. More specifically, by subclassing *PreferredResource* the importing ontology defines for which type of resources (i.e., individuals) the user can define preferences. Similarly, subclasses of *ContextObject* should be defined in order to indicate which type of individuals the *Context* consists of (i.e., the business process, activity, task, and the user).

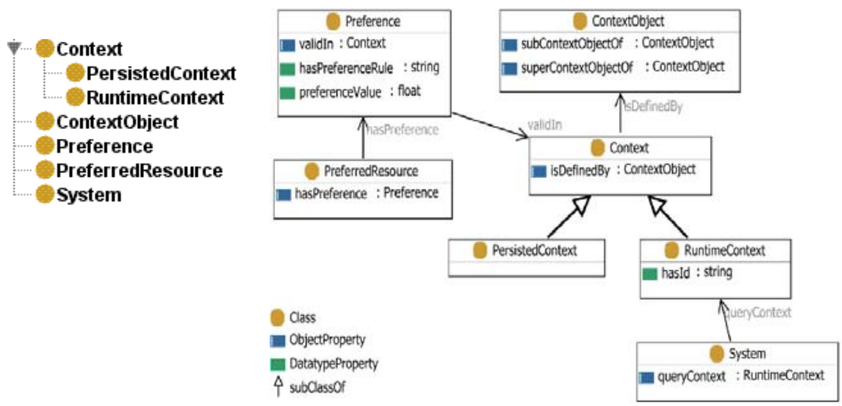


Fig. 6. Preference Ontology, Class hierarchy (left), Class diagram (right)

Furthermore, we differentiate between a *RuntimeContext* and a *PersistedContext*. The RuntimeContext reflects the user's current context and changes dynamically with the user's interactions within the system. This context may be used to track user's behaviour in the system. However, if a user's interaction is logged in the system as a persistent activity (e.g., the creation of a new document) the user's current context will be persisted (using the *PersistedContext*).

The RuntimeContext and PersistedContex are utilised by the Context Observer to extract the current business context, and hence, enable resource delivery based on that business context.

A preference is an n-ary relation between a user, multiple resources, and a preference value.

Each preference (i.e., n-ary relation) is expressed as a logical rule, represented in SWRL.

Utilising logical rules, for expressing context-sensitive user preferences, our EAMS features a flexible preference model. One rule is used to assign different preference values to different information resources based on relevant criteria of a particular user. Therefore every information resource may be assigned with different preference values by different preference rules (i.e., by different users and/or business contexts). Another flexibility of the proposed preference model comes from an implicit representation of preferences. Since preference values are not pre-computed and persisted in the system, just adding one preference rule may significantly influence the whole preference model. Also adding a common preference to the preference model (i.e., a preference valid for all users) may be as easy as adding only one preference rule. Moreover updating existing resources, or adding new ones, does not mess up all previously created preference values. In this way, a user is given a great freedom to create particular preferences for particular processes, activities, tasks, and even to aggregate multiple preference values[3] for one resource into a final score.

Preference rules, created by the editor, are serialised to its SWRL representation and stored in the preference ontology. Finally, preference rules may also be removed (or updated) using the Preference Editor.

Rule-based, formal, reasoning plays a central role for managing user's attention in the EAMS. As we already explained in the previous section, a preference is expressed as a logical rule, represented in SWRL. The rule matches an information resource with the desired criteria (i.e., particular business context etc.) and assigns a preference value and linked actions to them. Preference rules are stored in the knowledge base, and they are evaluated by the reasoner when certain events occur within the system. Events are triggered whenever a user changes her/his run-time context. A user may utilise these events to create ECA rules.

Currently we allow user to use preference rules as the action part of an ECA rule, and do not explicitly define the condition part[4]. Since a few preference rules may be defined for one particular event, our EAMS may be seen as a reactive system capable of executing basic form of complex actions (i.e., actions that comprise of simple or atomic actions). Enhancing the system with more complex operators, for combining atomic action into complex actions, is a subject of our future work.

It is a task of the reasoner to react on events and evaluate appropriate ECA rules. The evaluation procedure is as follows: on an event, the system issues a SPARQL query which starts the reasoner engine. The reasoner takes into account a relevant preference rule (with the respect to the SPARQL query) and evaluates the rule. In cases there are more than one preference rules defined for a particular information resource, the reasoner applies an aggregation function combining the single preference values into an aggregated score.

Once the rules are positively evaluated, a notification is generated and provided to the user both as a web bulletin and an RSS feed. The RSS feed points to a web page containing the results of the evaluated rules. By adopting RSS as a notification mechanism instead of sending emails, we allow the user to select whether s/he wants to be notified immediately about relevant information, periodically or when s/he wants to be notified. Most RSS readers allow the user to specify when and how they want to be notified abut updates in the feeds they receive.

[3] Preference aggregation is computed in cases where there exist more than one preference rule for an information resource.

[4] The condition, if required, may be formally expressed in the action part (i.e., in the preference rule).

We use KAON2 as an underlying inference engine for managing, querying and reasoning about the ontologies. KAON2 supports the SHIQ(D) subset of OWL-DL, and DL-safe subset of SWRL rules. Also a good portion of (but not entire) SPARQL specification is supported, and a number of built-in predicates have been implemented in this reasoner.

4 Pertinent Technologies and Related Work

Pertinent technologies and related work includes research in attention aware systems, just in time and context-based information retrieval and delivery systems, web monitoring and alert systems and RSS-oriented enterprise systems. There has been considerable research done on attention aware systems that address the information overload problem (see for example the Computer in Human Behavior special issue on attention aware systems [13]). According to [10] attention-aware or attentive systems are software systems that support users' information needs by keeping track of what users are writing, reading, looking at or talking to and suggesting information that might have beneficial influence to them. Just-In-Time Information Retrieval agents (JITIRs), is a class of software agents that search and present relevant to the user's local context accessible information, without any user involvement, in a way that it does not distract him from his current task [12]. Web monitoring and alert systems are trying to detect and deliver any change concerning Web pages or other information that travels on the Web to any interest party. Finally, we present pure RSS-oriented systems that are enterprise-wide web feeds platforms that enable individuals, workgroups, and the entire organisation to subscribe and receive information either from external or internal blogs, wikis, websites or other source. RSS (Really Simple Syndication) is a communication technology that allows subscribers to automatically receive information on any topic as soon as it is published over the Internet. With an RSS reader (aggregator) like FeedReader or NewsReader, a user can establish a direct link to an information source and automate the watching process for updates reducing loss of time from searching and fetching again and again. This new approach increases the opportunity for knowledge acquisition, sharing and collaboration among knowledge workers streamlining internal and external communications.

4.1 Attention Aware Systems

Researchers from IBM and MIT designed and developed the *Simple User Interest Tracker* or *SUITOR* [10] which is a framework for developing attentive information systems that track computer users through multiple channels to determine interests and to try to satisfy information needs. SUITOR implements four main modules: a) watching user's actions to infer user's current mental state and needs, b) storing user's actions to create and maintain user's model, c) searching information from the digital world and scanning user's local hard disk and d) ranking and suggesting relevant information sources through peripheral display. Having enough input information from these modules, SUITOR can infer user's current interests and propose relevant information sources from local and remote databases that have previously gathered and stored.

The *Attentional User Interface project* [8] developed methods for inferring attention from multiple streams of information, and for leveraging these inferences in decision-making under uncertainty. These methods have been used in illustrative applications of the use of attentional models. Applications focus on the design of new interfaces that take into account visual attention, gestures and ambient sounds as clues about a user's attention. These clues can be detected through cameras, accelerometers and microphones or other perceptual sensors and, along with the user's calendar, current software interaction and data about the history of user's interests, they provide valuable information about the status of a user's attention. The same project built Bayesian models aiming at dealing with uncertainty and reasoning about current or future user's attention taking as input all of the above clues. Moreover, it introduced the economic model of attention and information. The model computes the expected cost of disruption user's current activities and infers whether and how to alert the user and display incoming messages.

A significant stream of related work deals with the design of attentive user interfaces (e.g. see [15], [16]). Shell et al [14] identify five key properties of attentive user interfaces: sensing attention, reasoning about attention, communication of attention, gradual negotiations of turns, and augmentation of focus. Wood et al. [17] make the distinction between visual and auditory attention and discuss five themes that concern the nature and measurement of visual attention.

In order to analyse the issues related to the design of attention aware systems, Roda and Thomas [13] have identified three aspects of attention management: (1) Detection of current user's attentional state. The system needs to establish what are the user's goals and current tasks, where is the user's attention focussed, and what is happening in the environment, (2) Detection and evaluation of possible alternative attentional state. The system establishes whether alternative foci are available, how important they may be for the user, and the cost effectiveness of possible focus switches, (3) Strategies for presentation of alternative states to the user (or maintenance of current focus). The system defines the strategies best suited to present the user with alternative foci.

In comparison to attention aware systems, our system does not include sensor-based mechanisms for detecting the user's environment. We argue that for enterprise attention management, non-sensory based mechanisms provide a wealth of attentional cues such as users' scheduled activities (e.g. using online calendars), users' working context (e.g. by querying workflow or enterprise systems) and user's communication and collaboration patterns (e.g. using groupware and other communication tools). However, our system is tailored to support enterprise attention management by taking into account the business context and working preferences.

4.2 Just-in-Time Information Retrieval and Context-Based Information Delivery

In the domain of text editing, Rhodes developed three agents with attention management characteristics [12]. The *Remembrance Agent* (RA) is attentive because it monitors user input from several sources and displays relevant documents in a non-distracting manner. The RA agent is using a display buffer at the bottom of the Emacs window to present a list of related documents while the user is writing or reading his document using Emacs. Suppose we were writing a paper concerning "Attention Management" using Emacs. At the bottom of the Emacs window we could see the list of the related documents that came for example from the INSPEC database of conferences and journal abstracts in order of relevance. The order of relevance of the suggestions being presented is based on word co-occurrence, meaning that both our paper and the suggested citations contain the same words, for example the words "information retrieval", "monitoring system" and "attention management".

The *Margin Notes* (MN) agent automatically adds hyperlinks to related documents as the user browses a certain web page. While the web page is loading, MN searches in its database of pre-indexed email archives or other text files and if a relevant file is found then MN adds a black strip on the right margin presenting this annotation.

The third agent is *Jimminy* that detects the user's physical environment through sensors such as the user's physical location.

All three agents use Savant, one common information retrieval engine that produces the suggested documents. Savant receives the text query created by the agents, performs a keyword-based comparison with the pre-indexed files and produces a ranked list with the most similar documents.

Context-based, proactive delivery of information refers to delivering information to users based on context, e.g. activities, organisational role, and work outputs. Maus [11] and Abecker et al. [1] presented workflow management systems that recommend relevant information to users based on the process context. The Watson system [6] provides users with related documents based on users' job contents such as word-processing and Web browsing. Ahn et al. [3] provide a knowledge context model, which facilitates the use of contextual information in virtual collaborative work. In Agostini et al. [2], organisational context is modeled to provide awareness of other users' activities in a shared workspace. In our work, context-based delivery of information is coupled to attention aware delivery of information and is also used for triggering actions.

Besides not following an agent-based architecture, our EAMS differs in that it supports semantics-based information retrieval, allowing for retrieving content that contains semantically similar concepts.

4.3 Web Alert Systems

There are several commercial systems and few research prototypes developed towards monitoring web changes. *WebCQ* [9] is a server-based research prototype that enables users to detect and deliver information changes on the Web by making use of continual queries. Continual queries are standing queries that the system issue once and run constantly monitoring the set of Web sources for updates. The definition for each continual query is made by the query itself written in SQL, a trigger condition and a stop condition that both might be time-based or content-based. As soon as the trigger condition of the query becomes true the system pushes the new result to its relevant users. Before disseminating new results the system evaluates the stop condition in case of time expiration. The prototype utilizes four modules: 1) the change detection robot that controls the user selectable information that the system monitor, 2) the proxy cache service that aims to reduce the frequency of the requests to the remote servers, 3) the presentation tool that highlights the changes detected and 4) the notification service that controls when, how and what information should be presented to the user depending on his preferences for the time interval, through server-pushed means (emails or web pages) or client-pulled queries.

Our EAMS offers similar functionalities to Web alert systems because it can monitor web pages with textual content (html, php, asp etc) but also RSS feeds and file servers for possible changes. In this respect, our system provides wider source coverage than WebCQ and similar to the coverage provided by commercial tools such as *WebSite-Watcher* (www.aignes.com). Moreover, it can channel notifications triggered by Web changes through a reasoner that takes into account user preferences and notifies users accordingly.

4.4 RSS-Oriented Enterprise Systems

RSS technology has been adopted in enterprise systems as a means to provide instant information between enterprise communities, customers and partners. Enterprises can benefit from this mean of information delivery and improve awareness of business events and thus immediate response to new challenges and opportunities by leveraging the new lingua franca for exchanging data over the Web as well as helping them solve problems, improve decision-making and spark innovation in general.

Based on publicly available information, we outline two enterprise platforms that utilize RSS technology: Attensa and NewsGator. Each one of them provides an enterprise-wide Web feed environment aiming to manage information overload from everyday news, emails, and newsletters, to separate the relevant from the irrelevant content and distribute the right information to the right people at the right time.

Attensa (www.attensa.com) provides *Attensa Feed S*erver (AFS) an integrated solution aiming at securely handling RSS feeds. AFS utilizes RSS, ATOM & XML to share information to interest parties as soon as it is published. AFS is designed to connect employees with the external and the internal world of an enterprise environment. It is installed behind the corporate firewall and provides a means to create a secure, scalable enterprise RSS environment providing up-to-date targeted information. Today online publishing tools, RSS feeds together with blogs and wikis inspire collaboration and spawn communities and social networks reinforcing group memory and collaborative filtering. This is the idea behind the Attensa philosophy that creates an RSS network whether employees working on a desktop or a mobile phone or a PDA. Moreover, RSS feeds can be read in different RSS readers through Attensa's AttentionStream synchronization. This technology ties the server to the RSS network and therefore RSS readers, so feeds can be read, continuously updated, filed and deleted regardless of when and where they are accessed. Moreover, AttensionStream has analytics and reporting capabilities that profile user behavior to predict and identify the most effective communications channels between users-employees. User behavior can be tracked down by assessing the time certain users read certain RSS feeds, which feeds they are going to firstly, which of them are ignored or deleted or even which are forwarding to which others predicting potential common user interests in certain user groups. Reports can be used to identify must-read feeds in order to optimize communication channels.

NewsGator Enterprise Server (NGES) (www.newsgator.com) is a centrally managed and administered RSS aggregation platform. It offers access to RSS and Atom sources, search tools to find relevant feeds, multiple user options for feed reading and enables users to collaborate and discuss important topics. Concerning common problems with bandwidth consumption from multiple users that subscribe to the same feeds, NGES can get the feed one time and distribute it internally to all subscribed users. Besides pure RSS feeds, NGES can aggregate and deliver content to its subscribed users from internal portals and projects, enterprise applications and databases (CRM, ERP, HR), premium content providers such as Thomson or Factiva, blogs, wikis, e-mail applications or even relevant information from audio and video pod casts.

Centralized management and administration features of NGES include system and group administration, feed locking to prevent users from unsubscribing feeds, customizable taxonomy for easy feed discovery and credential management with automatic authentication for secure feeds. NGES supports collaboration between collaboration platforms like blogs and wikis notifying users that might not normally visit these tools but need to stay aware, with interesting content that traffic inside them. Employees can automatically be subscribed to feeds in from blogs and wikis by embedding one-click “chiclets” for their interests reducing frequent checks for updated content. Other collaboration features include clippings, tagging, recommended feeds, discussions and email feeds. Any user can clip or drop an interesting item in a folder, add a comment and let others subscribe to this item, tag an article or categorize it for easy sorting, get some recommended feeds based on viewing habits of others, start a discussion about some interesting topic, or even turn any email into an RSS feed.

Similarly to enterprise RSS systems, our EAMS provides RSS aggregation facilities but it also supports personalised user notifications and proper reaction on notifications by taking into account user preferences.

4.5 Other Technologies — AttentionXML

So far we have presented platforms and applications that facilitate the collection of data, relevant to the user’s working context. In other words, these efforts are trying to monitor and analyze how the user interacts with his/her environment and conclude his/her interests and habits or where they pay attention to.

One way to track users’ attention while they are working inside web applications such as web sites, blogs and wikis is to use AttentionXML [5], an open specification that enables tracking what users are reading, looking at or listening to and collecting usage data in order to predict usage patterns and feed recommendation systems. More specifically these attention metadata record what sites, blogs and feeds the users read, when they did it for the last time, how long they spend on them and other information. In Attention.XML each blog, feed, etc. is represented in a hierarchical outline structure which is annotated with per-feed and per-post information which captures such information as, the last time the feed/post was accessed, the duration of time spent on the feed/post, recent times of feed/post access, user set (dis)approval of posts, etc. Nevertheless, this standard provides only the basic information of user’s attention, e.g. does not capture how the user interacts with the content or what the user likes and dislikes.

5 Conclusions

In this chapter we presented a novel approach for managing attention in an enterprise context by realising the idea of having a reactive system that manages not only alerting a user that something has been changed, but also supporting the user to react properly on that change. In a nutshell, the corresponding system is an ontology-based platform that logs changes in internal and external information sources, observes user context and evaluates user attentional preferences represented in the form of ECA rules.

Since the presented system is currently under deployment in a real-environment, results from a formal evaluation are still missing. However, initial assessments have shown that the users find the system useful and usable. Additionally, efforts to setting-up the whole system are limited, which is one of important characteristics of semantics-based applications. A formal evaluation of the system will follow; in this we will evaluate the system's ability to provide relevant recommendations (precision and recall metrics) as well as the overall added value to the users (in terms of e.g., reduction in time to perform a task) contrasted to the effort needed to setup and use the system.

Future work will be toward further refinement of ECA rules for preference description and automatic learning of preferences from usage data using and machine learning techniques. Considering the existence of hierarchical relations that exist in the Log Ontology which models interactions (events) that happen in the EAMS, we can further utilize Generalized Association Rules in order to mine and discover interesting patterns in the system's usage. Generalized Association Rules improve upon standard association rules taking into account a taxonomy. In particular, let I = $\{i_1, i_2, \ldots, i_m\}$ be a set of items. A GAR is a deduction of the form:

X => Y, where $X, Y \subset I, and \cdot X \cup I = \varnothing$ and no item in Y is an ancestor of any item in X.

For example, GARs can help discovering that many users involved in business process X are referring to documents created by employees of department Y or that they often consult the same employees by exchanging forum messages.

Moreover, future work will tackle the problem of the automatic definition of events. Indeed, in the current version events are defined in advance and “hard-coded” in the process workflow, which supports execution of events, but decreases the adaptivity of the system. We will work on methods for event discovery from log data, which will enable dynamic changes in the configuration of the system.

Resources

A Key Books

Allport, A. (1992) Attention and Performance XIV, Attention and Control: Have We Been Asking the Wrong Questions?, In D.E. Meyer & S. Kornblum (Eds.), London, MIT Press.

Davenport, T., Beck, J. (2001) The Attention Economy: Understanding the New Currency of Business. Harvard Business School Press.

B Key Journals

Communications of the ACM (ACM).

Computers in Human Behavior (Elsevier).

Data & Knowledge Engineering (Elsevier).

Decision Support Systems (Elsevier).

Expert Systems with Applications (Elsevier).

Knowledge-Based Systems (Elsevier).

IBM Systems Journal.

IEEE Transactions on Knowledge and Data Engineering (IEEE).

C Key International Conferences/Workshops

CHI – Computer/Human Interaction Conference.

COOP – International Conference on the Design of Cooperative Systems.

IUI – Intelligent User Interfaces International Conference.

D Software

Google Alerts are email updates of the latest relevant Google results (web, news, etc.) based on your choice of query or topic *http://www.google.com/alerts*

Nutch
http://lucene.apache.org/nutch/

Java 2 Platform, Enterprise Edition (J2EE)
http://java.sun.com/javaee/

Feedreader is a free RSS news aggregation solution
http://www.feedreader.com/

Newsreaders are programs that let you read newsgroups
http://www.newsreaders.com/

WebSite-Watcher
(www.aignes.com)

Attensa
www.attensa.com

NewsGator Enterprise Server
www.newsgator.com

References

1. Abecker, A., Bernardi, A., Hinkelmann, K., Kühn, O., Sintek, M.: Context-aware, proactive delivery of task-specific information: the Know-More Project. Information Systems Frontiers 2, 253–276 (2000)
2. Agostini, A., De Michelis, G., Grasso, M.A., Prinz, W., Syri, A.: Contexts, work processes and workspaces. In: Proceedings of the International Workshop on the Design of Cooperative Systems (COOP 1995), INRIA, Antibes, France, pp. 219–238 (1995)
3. Ahn, H.J., Lee, H.J., Cho, K., Park, S.J.: Utilizing knowledge context in virtual collaborative work. Decision Support Systems 39(4), 563–582 (2005)
4. Allport, A.: Attention and Control: Have We Been Asking the Wrong Questions? A Critical Review of Twenty-Five Years. In: Meyer, D.E., Kornblum, S. (eds.) Attention and Performance, vol. XIV. MIT Press, London (1992)
5. AttentionXML specifications, `http://developers.technorati.com/wiki/attentionxml`
6. Budzik, J., Hammond, K.W.: Anticipating and contextualizing information needs. In: Proceedings of the Sixty-Second Annual Meeting of the American Society for Information Science (1999)
7. Davenport, T., Beck, J.: The Attention Economy: Understanding the New Currency of Business. Harvard Business School Press, Boston (2001)
8. Horvitz, E., Kadie, C.M., Paek, T., Hovel, D.: Models of attention in computing and communication: From principles to applications. Communications of the ACM 46(3), 52–59 (2003)
9. Liu, L., Pu, C., Tang, W.: WebCQ - Detecting and Delivering Information Changes on the Web. In: Proceedings of the IEEE Conference on Information and Knowledge Management, Washington, DC, pp. 512–519 (2000)
10. Maglio, P.P., Campbell, C.S., Barrett, R., Selker, T.: An architecture for developing attentive information systems. In: Knowledge-Based Systems, vol. 14, pp. 103–110. Elsevier Science, Amsterdam (2001)
11. Maus, H.: Workflow context as a means for intelligent information support. In: Akman, V., et al. (eds.) CONTEXT 2001. LNCS (LNAI), vol. 2116, pp. 261–274. Springer, Heidelberg (2001)
12. Rhodes, B.: Just-In-Time Information Retrieval. Ph.D. dissertation, MIT Media Arts and Sciences, Cambridge, MA (June 2000)
13. Roda, C., Thomas, J.: Attention Aware Systems: Theory, Application, and Research Agenda. Computers in Human Behaviour 22, 557–587 (2006)
14. Shell, J.S., Selker, T., Vertegaal, R.: Interacting with groups of computers. Communications of the ACM 46(43), 40–46 (2003)
15. Vertegaal, R.: Attentive User Interfaces, Introduction to Special Issue. Communications of the ACM 46(3), 30–33 (2003)
16. Vertegaal, R., Shell, J.S., Chen, D., Mamuji, A.: Designing for augmented attention: Towards a framework for attentive user interfaces. Computers in Human Behaviour 22, 771–789 (2006)
17. Wood, S., Cox, R., Cheng, P.: Attention Design: Eight Issues to Consider. Computers in Human Behaviour 22, 588–602 (2006)

Chapter 13

e-Welfare as a Client Driven Service Concept

Katariina Raij and Paula Lehto

Laurea University of Applied Sciences

Abstract. This paper describes the processes which made the development of e-welfare services with and for elderly people possible. This kind of development process is seen important because of the increasing number of elderly people and economical challenges, a welfare service sector has to face. We believe that by applying technology new solutions can be found. It means that different experts learn to work together and it also demands new kinds of environments. We also see that students in professional higher education should be involved from the beginning in the development of new kinds of technology based welfare services, service concepts and client driven working models. We, at first, introduce the model of the new kind of environment as a meeting place and then proceed to describe an action model in which an authentic research and development project forms a learning environment, and in which all the participants are seen as equal partners and learners.

The research and development project, Going Home, will be used as an example of how technology based e-welfare services can be designed with and for elderly people. In the Going Home project the two channel interactive Caring TV forms a technical platform for e-welfare services. The research findings show of how elderly people become more active and responsible for their own welfare status. Virtual guidance and counselling services according to the conceptions of elderly people are identified and the client driven programme production process is described. At the end the indicators of quality of life are presented based on elderly people's own descriptions.

Keywords: e-welfare services, a client driven service concept, a centre of welfare competence, Learning by Developing action model, Caring TV, the indicators of quality of life.

1 Background

1.1 The Challenge in Elderly – The Demand of New Kind of Competence Sharing

The population is expected to grow in all major areas of Finland (Jokiranta 2006). The highest growth will take place in the 65—74 age group. This will create a significant increase in the proportion of senior citizens in the population. It also means an increase in costs in the health and social sector (e.g. Aromaa & Koskinen 2002). As the population ages, the challenges of elderly and geriatric care are a central area of development for private and public social service and health providers. Some

G.A. Tsihrintzis and L.C. Jain (Eds.): Multimedia Services in Intel. Environ., SIST 3, pp. 283–298.
springerlink.com

estimates indicate that more than 80% of daily services needed by senior citizens already come from outside the public sector.

In future, there will be more and more healthy, capable senior citizens. Solutions for them may consist of welfare technology applications related, for instance, to accommodation, daily life, information and service development. Promoting the well-being and resources of the elderly requires the competence of a multidisciplinary team of experts, whose work is based on the expectations and situations of clients and their friends and relatives. Increasingly, senior citizens want to and are able to live at home for as long as possible, despite chronic illness or decreased functional ability, which means that more tailored welfare services are needed in the home. The availability and compatibility of such services contribute to coping at home. More efficient productivity presupposes innovative technological applications about which unanimity prevails among innovative service designers. As Hyppönen (2004) has pointed out, the research and development of technology based service concepts are prioritized in welfare sector.

The Information Technology offers many possibilities in the field of welfare. The challenge is in creating possibilities for technology and welfare experts to share their knowledge and competences in order to arise and improve new ideas for the development of service innovations. We will use the Well Life Center (WLC) as an example of a new kind of environment built for the integration of welfare, technology and business competences.

The WLC has been built based on theories of how new professional knowledge can be created.

Professional expertise is found to be built of knowing (evidence based knowledge), understanding, and doing and situation management. The orientations of professional competence building are identified as doer's, client's, researcher's and working processes' orientations. The concept orientation, in this context, is defined as the way a person perceives the phenomenon of learning in the meaningful way (Boekaerts 1996). The types of knowledge in professional competence building, in turn, are identified as 1) theoretical knowledge, 2) knowledge embedded in skills and abilities, 3) moral knowledge and 4) experiential knowledge. (Raij 2000, 2003.)

In the studies of knowledge creation within an enterprise, Nonaka has introduced the idea of Ba, which is a meeting place as a physical, spiritual and virtual place. In meeting middle-up-down people, in an enterprise, have a possibility to explicit tacit knowledge by sharing their experiences, finding new, improving and duplicating. Interaction leads to create new knowledge. (e.g. Krogh, Ichijo & Nonaka 2000.) By applying the idea of a meeting place to the research findings, mentioned above, it was possible to construct the theoretical frame for the environment where technology based innovations are developed for and with clients. One of the questions which this raised was who were supposed to meet if we wish to create new professional, welfare related knowledge and new technology based service innovations? The answer was found from the orientations of professional competence building as welfare and technology doers, researchers, clients and service processes. In this way the types of knowledge in professional competence building are also present. Participants bring their own clients as well as their activities with services to be present in the meeting place. A value base was identified as aiming at a human being's good. (Figure 1).

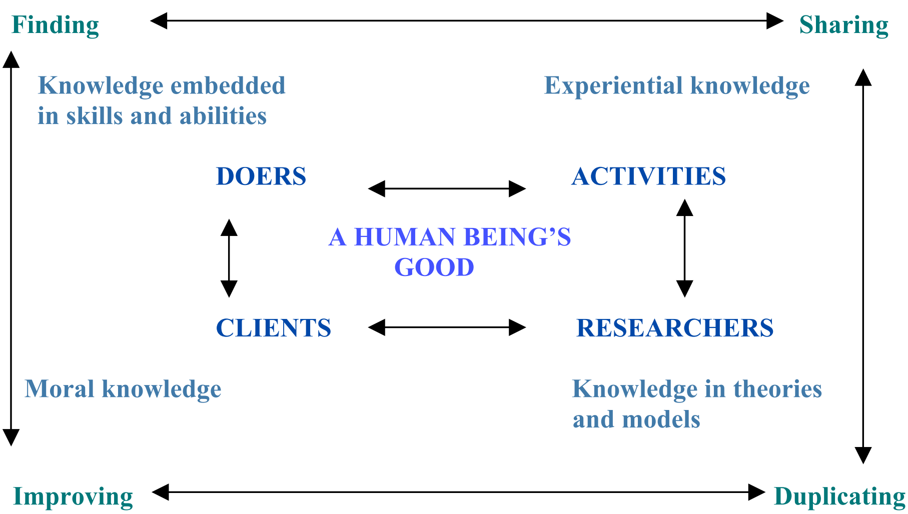


Fig. 1. The framework of the innovative environment where doers, clients, activities, and researchers (Raij 2000) create new professional knowledge by sharing, finding, improving, and duplicating (appl. Nonaka 2000)

Meanwhile Espoo City (the second largest city in Finland) presented its own strategies where significant expectations of the integration between welfare, technology and business were strongly expressed. When the proper building also was found, it was possible to begin constructing an innovative environment for the creation of new, welfare related and technology based professional knowledge. The latter are identified as innovations (products, working models and working culture) where welfare knowledge, technology and business are integrated. The value base, a human being's good, provided a guide for formulating the name "Well Life Center" in 2004. By 2005 all the participants representing higher educational institutes as well as public -, private - and the third sector organizations with welfare -, technology - and business competences were found and located within the WLC.

The WLC is coordinated and run by the Laurea University of Applied Sciences. The tasks described in the Finnish law (351/ 2003) include: higher education that responds to the world of work and its developmental needs, and is based on research and artistic principles; applied research and development work that fosters regional development and takes into account the industrial structures of the surrounding areas; and support for individual professional growth. According to these tasks, universities of applied sciences are closely linked to the world of work and therefore rooted in praxis. Scientific demands are related to pedagogy as well as to research that develops the working life and supports regional development. Demands for professionalism are evident in the idea of professional expertise based on the world of work and its development needs, and in the requirement for individual professional growth. The WLC is, as mentioned above, based on the co-operation and partnership between universities, public -, private - and the third sector organizations. Organizations operating in active partner network are creating benefits for themselves with the help of Laurea's long-term developmental partnership. This developmental partnership is the core added value for all involved organizations.

1.2 Caring TV – As a Platform

Caring TV is one of the promising innovations discovered and developed in the WLC. It is shortly described below as an example of evidences in the WLC. Caring TV is developed by the Laurea University of Applied Sciences, TDC Song and Espoo City. It is a two channel interactive TV through which guidance and support services will be given as far as various programmes to improve and promote the capacities of elderly people living at home. The content of guidance and support services are planned together with clients and with the supervision of experts. In planning these services an elderly person is taken into account as an active partner and as a holistic being (e.g. Rauhala 1995) with his or her own knowledge base, skills and abilities, and values and experiences (e.g. Raij 2003). This we call a client centred method. In the first phase a municipality buys TV channel to selected elderly people. In the second phase it will be offered to the private sector. This will mean that everyone living at home can buy a product which includes both the technology and content production (Piirainen & Raij 2006, Raij 2008).

Laurea is responsible for the research and development of the Caring TV – concept and content production, while TDC Song as a private company provides the technology and Espoo City the guidance and support services. Other private companies and municipalities are also included (e.g. Vantaa), and experts from the third sector have been invited to the project. Private companies, e.g. Medixine Oy as the developer of the tools for advanced e-services, PhysioSporttis Oy and Lääkärikeskusyhtymä Oy as the developers of physiotherapeutic and medical e-services are included, as well as experts from the third sector. The development of Caring TV, as a WLC product, opens new doors and gives us valuable knowledge on how to proceed towards the development of a virtual service clinic. It also has taught us how to introduce a new technological innovation to an end user by proceeding from a user centric to a user driven action model. This has been made possible by integrating the expertise offered by a university of applied sciences, a private enterprise and a public sector and by listening to the clients as end users. It means integrating welfare expertise, technology -, business - and research and development expertise in the WLC.

E-welfare services have been developed in three research and development projects, Coping at Home and HOME, funded by TEKES/ Finn Well programme and EU/ InnoElli Senior programme which aim to find new solutions for elderly people living at home and for municipalities dealing with current problems in health and social services. In the future project, Safe Home, funded by EU/ EAKR programme, elderly people, suffering from dementia, will be included to the development of virtual services.

In the Coping at Home project (e.g. Piirainen & Raij 2006, Piirainen & Sarekoski 2008) our pilot group consists of 25 family care givers living in Espoo, whereas in the HOME project (Lehto 2008, Raij 2008) there are 60 high risk clients from Vantaa City and 40 elderly people using services delivered by special service houses in Espoo, Turku and Lappeenranta. The interest was in discovering new, technology based solutions which support elderly people in staying at home and improving their quality of life by allowing them to have more control of their own lives. Laurea is responsible for the research and development of Caring TV – concept and participative content production, while Videra Oy as a private company provides the technology. The development of Caring TV has opened new doors and given us valuable knowledge on how to proceed towards the development of e-welfare service concept. Caring TV also offers a learning environment for students where they can achieve new competences by working together with educators, working life experts and clients, according to the Learning by Developing action model developed at Laurea (e.g. Raij 2007).

2 Action Research Enabling the Development of a Client Driven Service Concept

Action research has its roots in pragmatism and social sciences. The challenge has been in integrating the concepts of action and research as e.g. Cohen & Manion (1980) have pointed out.

Heikkinen (2006) divides action research into different elements by scope: individual, group, intergroup relations, organisation and regional network. According to Kuula (1999) the basic principles of action research are practice orientation, participation and the creation of new activities, interventions or knowledge related to change. Informants in action research are active participants in the research process. The basis of the relationship between researcher and participants are cooperation and participation. According to Creswell (2005) the criteria of the participative action research are fulfilled when clients, as end – users, are actively involved. We call this a client driven method in our studies.

The aim of action research is based on practice and is seeking answers and solutions to problems or settled tasks. The research process is circular and often begins with data collection in order to find and clarify questions or problems that can be solved and developed. The research process continues with planning, implementing and evaluating together with participants. The data collection happens in every phase of the action research. Action research analyses the background of actions, reflects on and develops alternative solutions to problems in order to achieve aims, produces new understanding and operating models. (Aaltoila & Syrjälä 1999, Kuula 1999, Heikkinen 2001, Anttila 2005, Heikkinen et al 2006).

The knowledge in action research can be classified and considered using the knowledge interests proposed by Habermas (1974). The technical knowledge interest is examined using quantitative data acquired through customer surveys. The use of action research represents the emancipatory knowledge interest, which relates to consciously generating new knowledge. For example in action research approach the proactivity refers finding new kinds of solutions for the change in the process. The critical knowledge interest can be seen in multiple dimensions and involves rich methods.

Methodologically, the data collection process can be described from the point of view of the approach and the method. It is used to acquire comprehensive information on the subject of the research. Triangulation can be used to analyse the researcher, the data collection process, the information source, the analysis unit or the theoretical and methodological approaches. Triangulation related to the information source refers to the perspective of the time, place and people (cf. Begley 1996; Flick 1998; Silverman 2001; Anttila 2005). Triangulation is used to acquire comprehensive information on the subject or the interest of the research. Triangulation related to the information source refers to the perspective of the time, place and people (cf. Knafl & Gallo 1995, Begley 1996, Flick 1998, Silverman 2001, Polit & Beck 2008).

The case study method often represents a qualitative and quantitative research approach. Case studies can focus on the past, present or future. The case is selected in accordance with the purpose of the research and the nature of the research questions or task. The case can be an event, a phenomenon, an individual or a group, placed for example in a specific context. Case studies provide descriptive, explanatory or comparative data (Järvinen & Järvinen 2004, xxx).

3 Learning by Developing as an Action Model

Learning by Developing (LbD) as an action model is close to the pragmatic learning theories developed e.g. Dewey (1899 and 1934), Bereiter and Scardamalia (1993), Engeström (1987 and 2001) and Hakkarainen, Lonka & Lipponen (2004). The stages of Learning by Developing (Raij 2007) were identified at Laurea by examining the ideas of lecturers who were conducting research and development projects (n= 5+25), and, as the process went on, their changing conceptions regarding the processes of implementation of R&D projects. The Learning by Developing model centres around a development project that is genuinely rooted in the world of work, which requires research-oriented and developmental collaboration between lecturers, students and workplace experts. Its starting point is a vision of the three tasks of the university of applied sciences (education, regional development, and research and development) as an integrated whole. An R&D project forms a learning environment, where progress is made through the identified stages and the outcomes are learning in individuals and in a community, and the production of new knowledge. The new knowledge becomes evident as new products, services, processes and operating models, and as a renewed work culture. At a university of applied sciences, it is essential that the creation of new knowledge also become explicit in the form of skills in doing. The aim is to develop the kinds of competence that transcend the traditional dichotomy between the professional and the scientific. In this operating model, the basis of the production of new knowledge lies in cultural interaction and in sharing, transferring and transmitting competence. Aristotle's *theoria*, *poiesis* and *praxis* are also present in the definition of the model's various stages. Theoria is apparent in the presence of empirical data, concepts and new knowledge; poiesis in the technical skills acquired in workshops; and praxis in collaboration with clients that uses the technical skills that have been acquired, aimed at inspiring engagement and doing good (Figure 2).

Fig. 2. Stages of LbD as identified by Raij 2007

According to respondents, some of the defining characteristics of Learning by Developing that could be identified were authenticity, partnership, trust, creativity and research-orientation. Authenticity arises from the genuine workplace R&D projects that form the learning environment, where participants work to find new things and produce new knowledge. Partnership is built on trust and on a commitment-inspiring agreement. All partners participate as equals, sharing experiences and finding meanings in order to produce new knowledge in their varying roles and responsibilities. There is room for every partner's creativity, which also leaves room for professional growth. The production of new knowledge and the development of competence become evident as the work progresses. The LbD model allows continuous, systematic and practically oriented development work, during which methodological solutions and actions can flexibly be changed and improved. All the participants take an active part in the change and research process. Planning, action and evaluation alternate and repeat in parallel cycles.

4 The Processes Leading to the Development of a Client Driven Service Design Described as a Case Study

The process leading to the development of a client driven service design will be described and demonstrated through the example of the conducted action research study named as Going Home project. In this study participative programmes, guiding and supporting services and a client driven service concept were developed in the context of CaringTV.

The purpose of the GOING HOME project was to further develop, productise and expand the use of intelligent welfare services provided via Caring TV. The aim was to use the latest IT knowledge to produce new applications that are based on online interaction in order to support elderly people living at home. The project developed the model for e-health services by producing guidance, counselling and support services for elderly people. The services and their content were produced through participative, rehabilitating and health-promoting programmes using the latest welfare technology. The model is based on client driven approach where the most essential feature is clients' active participation in the research and development process (Lehto 2008).

4.1 Research Findings as the Basis of Programme Production

Action research and the LbD model were applied in the research and development process. The action research approach was adopted as the starting point for the project, which meant that each stage of the project aimed to proactively produce new knowledge and actions. The methodological triangulation and multiple research and development methods were used in the project. Data collection methods included individual and group interviews, surveys and documentary analysis. The participants in the project were elderly clients (N= 93), professionals from service houses and home care (N=36). Also students and senior lecturers and technology and business partners from the universities of applied sciences and enterprises were actively participating in the research and development process (Lehto 2008).

Based on the data collection during the first phase of the study, following problems that elderly living at home have, were identified as: stimulation (refreshing) of the mind, safety of the environment, safety at home, activation of participation, belonging and togetherness, daily living activities, rehabilitation and physics and caring and monitoring. The themes for programme production were based on the data and were named as

* PROMOTION OF MENTAL HEALTH
* SUPPORT OF DAILY FUNCTIONS
* PROMOTION OF REHABILITATION AND PHYSICS
* GUIDING AND SUPPORTING OF CARE
* ENCOURAGEMENT OF PARTICIPATION
* ENCOURAGEMENT OF BELONGING AND TOGETHERNESS
* PROMOTION OF SAFETY OF THE ENVIRONMENT

Caring TV, as a service platform for new virtual e-services, were developed during the second phase of the study. The following figure demonstrates the framework of virtual e-services.

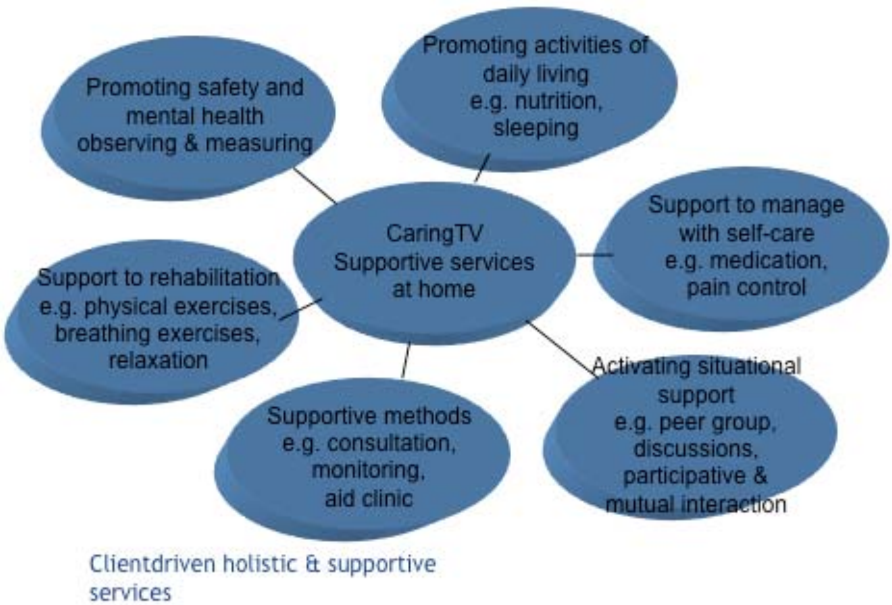


Fig. 3. Virtual guidance and counselling services by Lehto 2008

4.2 The Client Driven Programme Production Process

The purpose of Caring TV programme production in the GOING HOME project was to develop, create and produce new client-driven, innovative programmes and virtual welfare services. The main characteristics of the production initiative were to be client-driven, research-oriented and innovative. Systematic programme production and related research and development also promoted the commercialisation of the Caring TV concept.

Based on the above-mentioned study the objectives of the client driven programme production are; to develop and produce new client-driven programmes for the elderly, to test and produce programme contents and methods that support participation and to develop and produce content related virtual guidance and counselling services for the elderly.

Clients and experts participate actively in the CaringTV programme production process. Commitment and responsibility of the actors promote client driven approach. Research based orientation is a core in the process. Research orientation refers to data collection at various phases of the process and is systematic in order to specify the themes of the programmes. Also students as active actors are involved and participate in the programme production process. Their role is particularly important in identifying clients´ expectations and in planning, producing and evaluating programmes. Through participation students learn to integrate welfare technology and client-driven programme production in CaringTV. Applying the client driven approach in the programme production process, students' technological competence and competence in producing virtual interactive e-health services deepen. In future, as professionals, they are able to produce more new, interactive contents and methods for client encounters.

The programme production process can be described as follows:

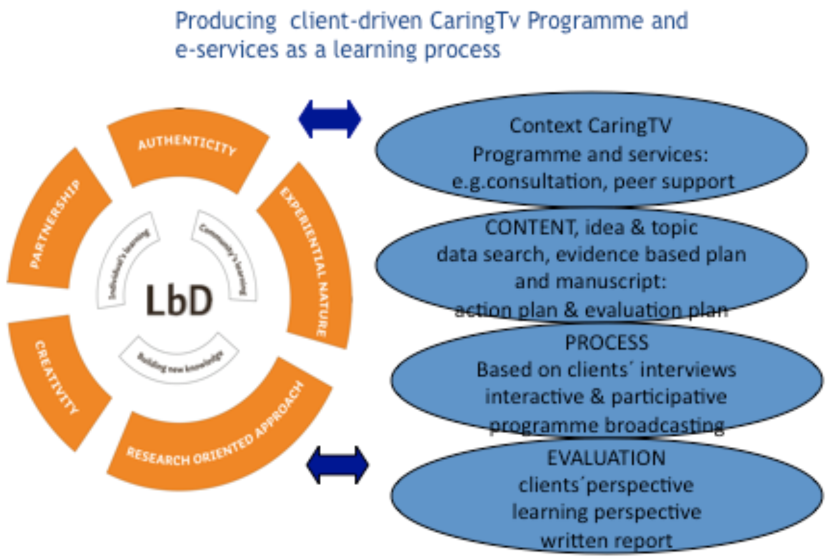


Fig. 4. The programme production process by Lehto 2007

The programme production process begins with the data collection in order to get the knowledge of clients´ needs and expectations. Based on this knowledge the idea and the topic of the programme will be chosen. The planning process continues by data search of the topic and main concepts. Literature search gives grounds and evidence for the content planning. Written manuscript includes contents, methods, supportive material and evaluation plan. The evaluation is ongoing process before, during and after programme production. The produced CaringTV programmes can be classified as informative or participative programmes and as virtual services.

For example in the GOING HOME project during two years 731 interactive and online programmes have been broadcasted. 43% (314) were classified as informative programmes and 39 % (285) were participative programmes. The virtual services included peer support, discussions, and services such as expert consultations, supported self-care and coping at home. The latter accounted for 10% (73). The rest of the programmes (53) varied in content and focused mainly on togetherness and belonging between clients (Lehto 2008).

5 e-Welfare Services Supporting Quality of Life in Elderly

The concept of quality of life offers many challenges for researchers (e.g. Rapley 2003, Walker 2005). In the definition given by the WHO Quality of Life Group (1998) quality of life is defined "as individuals' perceptions of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns". It is emphasized that "quality of life cannot simply be equated with the terms of health status, life satisfaction, mental state or well-being". Close to the WHO definition is the description presented by Felce and Perry (1993) where quality of life is defined " as an overall general well-being which comprises objective descriptors and the subjective evaluations of physical, material, social and emotional well-being together with the extent of personal development and purposeful activity all weighted by a personal set of values". In these definitions, there is a multidimensional concept related to an individual's perceived quality of life, which includes their own evaluations, but as Rapley (2003, 50) points out, it is unclear how an individual's perception of their position in life or a subjective evaluation is understood in relation to an individuals' mental state. In four of the five European countries, in which we have compared our concepts of the quality of life, the substantial experiences in measuring quality of life in old age are included. The main results show that there is a similar agreement in all the countries that health is the major determinant of quality of life. Good health seems to be one of the most important life goals according to the elderly people informants in the studies. This is in line with the value of quality of health and social services as a part of quality of life. According to ETAN (1998), information and communication technology seems to have the potential to improve the quality of life of older people but it also seems that elderly people form a disadvantaged group in the information society, and that they are not involved in product design. We made the same kind of conclusions but also found that the barriers are to be won.

The indicators of quality of life based on elderly peoples' own conceptions were identified in our studies in the following way (Figure 5).

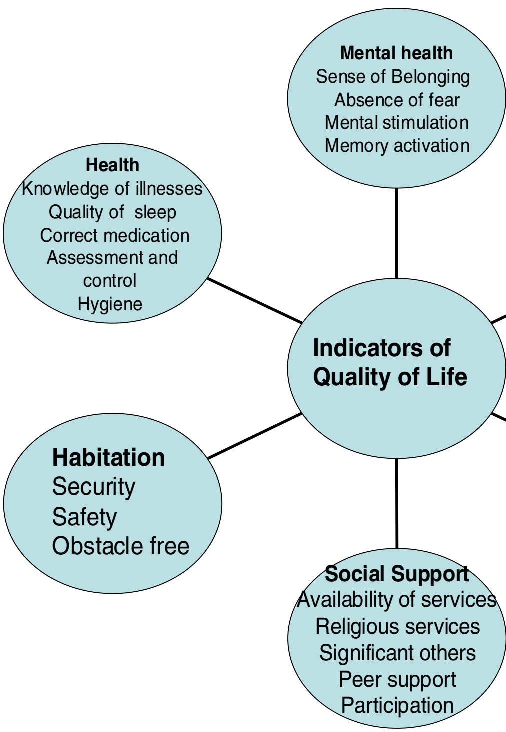


Fig. 5.

Benefiting the indicators of quality of life described by our clients, e- welfare services can be designed. The next challenge is the development of intelligent service design concept and intelligent e-welfare services and products that enhance the quality of life in elderly.

6 Conclusion

In this article we have followed the processes which have made the development of e-welfare services for elderly people possible. It has been seen important to take the challenges of the increasing number of elderly people into account. Another important issue has been the rapid development of information technology, which has opened a new door to the development of welfare services.

By introducing our Well Life Center as an innovative environment we have wanted to emphasize that new kinds of environments and new kind of network are needed for designing technology based welfare services. Experts representing different competences can learn a lot from each other, find new kinds of ideas and possibilities for new innovations. The only have to meet each other and work together in joint research and development projects.

Learning by developing as an action model has been described in order to show of how the creation of e-welfare services proceeds in cooperation between partners in the process, which also allows everyone to learn and share experiences. It also contains the cooperation with clients, who at the same time learn to take responsibilities for the development of their own services.

The process leading to the development of a client driven service design has been described by using the project Going Home as an example. It is supposed to show of how elderly people become more active and responsible for their own welfare status. In all the three projects it has been important to understand of how important it is to identify the meanings elderly people give to their experiences. By identifying the meanings it has been possible to develop also the guidance and support programme for elderly people of how to adapt to the new innovations.

At the end we present the indicators of quality of life as a challenge for our future work. They are valuable because they base on elderly people's own conceptions. They know best what is good for them. The development of e-welfare services will continue and designing with and for elderly people will be based on the indicators of quality of life.

Resources

A Key Books

1. 1.Aaltoila J. & Syrjälä L. 1999 Tiede, toiminta ja vaikuttaminen. In Heikkinen H., Huttunen R.&
2. P. Moilanen (eds.) Siinä tutkija missä tekijä – toimintatutkimuksen perusteita ja näköaloja, 11-24. Ateena Kustannus, Jyväskylä.
3. Anttila P. 2005 Ilmaisu, teos, tekeminen ja luova toiminta. Akatiimi, Hamina.

4. 4.Aromaa, A. and Koskinen, S. (ed.) (2002), Terveys ja toimintakyky Suomessa, (Health and functional capacity in Finland) Terveys 2000- tutkimuksen perustulokset, Kansanterveyslaitoksen julkaisuja B3, Helsinki.
5. 5.Bereiter C. & Scardamalia M. 1993 Surpassing Ourselves: An Inquiry into the Nature and Implications of Expertise. Open Court, Chicago.
6. 6.Boekaerts M. 1996 Coping with stress in childhood and adolescence. In Zeidner M. &
7. 7.Endler N.S. (Eds.) Handbook of coping, Theory, Research and Applications,452-484. John Wiley, NewYork.
8. Cohen L. & Manion L. 1980 Research methods in education. Biddles Limited, Guilford and Kings. Lynn.
9. Dewey J. 1933 The School and Society. University of Chicago Press, Chicago1899.
10. Dewey J. 1934 Art as Experience. Minton, Balch & Co., New York .
11. Dewey J. 1963 Experience and Education. Collier Books, New York 1963 (1938).
12. Engeström Y. 1987 Learning by Expanding. Orienta-Konsultit, Helsinki.
13. Engeström Y. 2001 Expansive learning at work: Toward an activity theoretical reconceptualization. Journal of Education and work 14 (1), 129 – 152.
14. Felce D. & Perry J. 1993 Quality of Life: A Contribution to its definition and Measurement. Mental Handicap in Wales – Applied Research Unit, University of Wales College of Medicine.
15. Flick U. 1998 An introduction to qualitative research. Sage Publications Ltd, London.
16. Habermas J. 1974 Knowledge and Human Interests. Boston. Beacon.
17. Heikkinen H., Kontinen T. & Häkkinen P. 2006 Toiminnan tutkimuksen suuntaukset. In Heikkinen H., Rovio E. & Syrjälä P. (ed.) Toiminnasta Tietoon, Dark Oy. Vantaa.
18. Heikkinen H. 2001 Toimintatutkimus – Toiminnan ja ajattelun taitoa. In Aaltoila J & Valli R. Ikkunoita tutkimusmetodeihin I. Metodin valinta ja aineiston keruu: virikkeitä aloitteleavalle tutkijalle. 170-185. PS-Kustannus, Jyväskylä.
19. Heikkinen H., Kontinen T. & Häkkinen P. 2006 Toiminnan tutkimuksen suuntaukset. In Heikkinen,H. Rovio E. & Syrjälä L. (ed.) Toiminnasta Tietoon, Dark Oy. Vantaa.
20. Järvinen P. & Järvinen A. 2004 Tutkimustyön metodeista. Opinpajan kirja, Tampere.
21. Knafl A.K. & Gallo A. M.1995 Triangulation in nursing research. In Talbot LA. (ed). Principals and practice of nursing research, 492-502. Mosby Year Book Inc, StLouis.
22. Kuula A. 1999 Toimintatutkimus. Kenttätyötä ja muutospyrkimyksiä. Vastapaino, Tampere.
23. Lehto P. 2008 b KOTIIN-hankkeen tutkimus ja kehittämistoiminta. In P.Lehto ed. KOTIIN-hanke.Asiakasvetoiset interaktiiviset ratkaisut ikääntyvän kotona selviytymisen tukena. 17-128. A63. Laurea-ammattikorkeakoulun julkaisusarja. Redfina Oy, Espoo.
24. Polit D.F. & Beck C.T. 2008 Nursing research. Generating and assessing evidence or nursing practice. Lippincott & Wilkins, Philadelphia.

25. Raij K. 2007 Learning by Developing, Laurea Publication A*58. Edita. Helsinki
26. Rauhala L. 1995 Tajunnan itsepuolustus. Yliopistopaino, Helsinki
27. Rapley M. 2003 Quality of Life Research. Sage Publications, London.
28. Silverman D. 2001 Doing qualitative research. A practical handbook. Sage Publications Ltd, London.
29. Hakkarainen L. , Lonka K. & Lipponen L. 2004 *Tutkiva oppiminen*. WSOY, Porvoo.
30. Walker A. 2005 Quality of Life in old age in Europe, In Walker A. Growing Older in Europe. Open University Press, McGraw-Hill House.
31. WHO 2008 World Health Organization 2001 International Classification of Functioning, Disability and Health, World Health Organization, Geneva.

B Key Survey/Review Articles

1. Lehto P. 2008 KOTIIN-hankkeen tutkimus ja kehittämistoiminta. In P.Lehto ed. KOTIIN-hanke. Asiakasvetoiset interaktiiviset ratkaisut ikääntyvän kotona selviytymisen tukena. 17-128. A63. Laurea-ammattikorkeakoulun julkaisusarja. Redfina Oy, Espoo.
2. Piirainen K. & Sarekoski I. 2008 Client-driven CaringTV concept for elderly family care givers living at home. Final report of the Coping at Home research. Laurea-publications A62. Vantaa, Edita priima Oy.
3. Raij K. 2008 HyvinvointiTV® tutkimus-ja kehittämishankkeena. In Lehto P. ed. KOTIIN-hanke.Asiakasvetoiset interaktiiviset ratkaisut ikääntyvän kotona selviytymisen tukena.129-146. A63. Laurea-ammattikorkeakoulun julkaisusarja. Redfina Oy, Espoo.

C Journals

Journal of Advanced Nursing
Journal of Education and work
Practice of Nursing research

D Key International Conferences/Workshops

1. The Third Sendai-Finland Seminar. Refurbishing the Elderly care –
2. New Health/Social Services and Network. 2007
3. The international conference Learning by Developing – New ways to Learn, Conference on Innovative Pedagogical Models in Higher Education, Laurea University of Applied Sciences. 2008

References

1. Jokiranta, T.: Espoon väestöennuste 2007-2016. Raportteja Espoosta 2/2006. Espoon kaupunki, Kehittämis- ja tutkimusryhmä. Espoo (2006)
2. Aromaa, A., Koskinen, S. (eds.): Terveys ja toimintakyky Suomessa (Health and functional capacity in Finland) Terveys 2000- tutkimuksen perustulokset, Kansanterveyslaitoksen julkaisuja B3, Helsinki (2002)

3. Hyppönen, H.: Tekniikka kehittyy, kehittyvätkö palvelut? Käyttäytymistieteellinentiedekunta, Helsingin yliopisto (2004)
4. Boekaerts, M.: Coping with stress in childhood and adolescence. In: Zeidner, M., Endler, N.S. (eds.) Handbook of coping, Theory, Research and Applications, pp. 452–484. John Wiley, NewYork (1996)
5. Raij, K.: Toward A Profession. Clinical leaning in a hospital environment as described by student nurses. An Academic Dissertation, University of Helsinki, Department of Education, Research Report no.166 (2000)
6. Raij, K.: Osaamisen tuottaminen ammattikorkeakoulun päämääränä. In: Kotila, H. (ed.) Ammattikorkeakoulupedagogiikka, Edita. Helsinki (2003)
7. Krogh, V., Ichijo, G., Nonaka, I.: Enabling Knowledge Creation: How to Unlock the Mystery of Tacit Knowledge and Release the Power of Innovation. University press, Oxford (2000)
8. Polytechnics Act 351/2003 Polytechnics Act (351/2003)
9. Rauhala, L.: Tajunnan itsepuolustus. Yliopistopaino, Helsinki (1995)
10. Piirainen, A., Raij, K.: Coping at Home, Annuals of the Kansei Fukushi Research Centre, Special Issue, Refurbishing the Elderly Care-evidence and Theoretical Targets, 13–21 (2006)
11. Raij, K.: HyvinvointiTV® tutkimus-ja kehittämishankkeena. In: Lehto, P. (ed.): KOTIIN-hanke. Asiakasvetoiset interaktiiviset ratkaisut ikääntyvän kotona selviytymisen tukena. 129–146. A63. Laurea-ammattikorkeakoulun julkaisusarja. Redfina Oy, Espoo (2008)
12. Piirainen, K., Sarekoski, I.: Client-driven CaringTV concept for elderly family care givers living at home. Final report of the Coping at Home research. Laurea-publications A62. Vantaa, Edita priima Oy (2008)
13. Lehto, P.: b KOTIIN-hankkeen tutkimus ja kehittämistoiminta. In: Lehto, P. (ed.): KOTIIN-hanke. Asiakasvetoiset interaktiiviset ratkaisut ikääntyvän kotona selviytymisen tukena. 17–128. A63. Laurea-ammattikorkeakoulun julkaisusarja. Redfina Oy, Espoo (2008)
14. Raij, K.: Learning by Developing, Laurea Publication A*58. Edita. Helsinki (2007)
15. Cohen, L., Manion, L.: Research methods in education. Biddles Limited, Guilford and Kings. Lynn (1980)
16. Kuula, A.: Toimintatutkimus. Kenttätyötä ja muutospyrkimyksiä. Vastapaino, Tampere (1999)
17. Creswell, J.: Educational Research: Planning, Conducting and Evaluating Quantitative and Qualitative Research. Merill, Upper Saddle River (2005)
18. Aaltoila, J., Syrjälä, L.: Tiede, toiminta ja vaikuttaminen. In: Heikkinen, H., Huttunen, R., Moilanen, P. (eds.) Siinä tutkija missä tekijä – toimintatutkimuksen perusteita ja näköaloja, pp. 11–24. Ateena Kustannus, Jyväskylä (1999)
19. Kuula, A.: Toimintatutkimus. Kenttätyötä ja muutospyrkimyksiä. Vastapaino, Tampere (1999)
20. Heikkinen, H.: Toimintatutkimus – Toiminnan ja ajattelun taitoa. In: Aaltoila, J., Valli, R. (eds.) Ikkunoita tutkimusmetodeihin I. Metodin valinta ja aineiston keruu: virikkeitä aloittelevalle tutkijalle, pp. 170–185. PS-Kustannus, Jyväskylä (2001)
21. Anttila, P.: Ilmaisu, teos, tekeminen ja luova toiminta. Akatiimi, Hamina (2005)
22. Habermas, J.: Knowledge and Human Interests, Boston, Beacon (1974)
23. Begley, C.M.: Using triangulation in nursing research. Journal of Advanced Nursing 24, 122–128 (1996)
24. Flick, U.: An introduction to qualitative research. Sage Publications Ltd., London (1998)

25. Silverman, D.: Doing qualitative research. A practical handbook. Sage Publications Ltd., London (2001)
26. Knafl, A.K., Gallo, A.M.: Triangulation in nursing research. In: Talbot, L.A. (ed.) Principals and practice of nursing research, pp. 492–502. Mosby Year Book Inc., St. Louis (1995)
27. Polit, D.F., Beck, C.T.: Nursing research. Generating and assessing evidence or nursing practice. Lippincott & Wilkins, Philadelphia (2008)
28. Järvinen, P., Järvinen, A.: Tutkimustyön metodeista. Opinpajan kirja, Tampere (2004)
29. Dewey, J.: 1933 The School and Society. University of Chicago Press, Chicago (1899)
30. Dewey, J.: 1934 Art as Experience. Minton, Balch & Co, New York (1934)
31. Bereiter, C., Scardamalia, M.: Surpassing Ourselves: An Inquiry into the Nature and Implications of Expertise. Open Court, Chicago (1993)
32. Engeström, Y.: Learning by Expanding. Orienta-Konsultit, Helsinki (1987)
33. Engeström, Y.: Expansive learning at work: Toward an activity theoretical reconceptualization. Journal of Education and work 14(1), 129–152 (2001)
34. Hakkarainen, L., Lonka, K., Lipponen, L.: Tutkiva oppiminen. WSOY, Porvoo (2004)
35. Lehto, P.: A CaringTV as a context for learning client-driven services. In: Presentation in The international conference Learning by Developing – New ways to Learn, Conference on Innovative Pedagogical Models in Higher Education, Laurea University of Applied Sciences (2008), `http://www.laurea.fi/net/en/06_Laurea/01_Communications/Seminars/LbD_conference_Feb_2008/03_Workshops/sh_3_2/LbDconference_3.2Lehto.pdf`
36. Lehto, P.: KOTIIN-hankkeen tutkimus ja kehittämistoiminta. In: Lehto, P. (ed.) KOTIIN-hanke. Asiakasvetoiset interaktiiviset ratkaisut ikääntyvän kotona selviytymisen tukena, vol. A63, pp. 17–128. Laurea-ammattikorkeakoulun julkaisusarja, Redfina Oy, Espoo (2008b)
37. Lehto, P.: Presentation in The Third Sendai-Finland Seminar. Refurbishing the Elderly care – New Health/Social Services and Network (2007)
38. Rapley, M.: Quality of Life Research. Sage Publications, London (2003)
39. Walker, A.: Quality of Life in old age in Europe. In: Walker, A. (ed.): Growing Older in Europe. Open University Press, McGraw-Hill House (2005)
40. WHO, World Health Organization, International Classification of Functioning, Disability and Health, World Health Organization, Geneva (2008)
41. Felce, D., Perry, J.: Quality of Life: A Contribution to its definition and Measurement. Mental Handicap in Wales – Applied Research Unit, University of Wales College of Medicine (1993)
42. Fernandez-Ballesteros, R., Zamarron, M., Ruiz, M.: The contribution of socio-demographic and psychosocial factors to life satisfaction, Ageing and Society. In: ETAN 1998 Ageing population and Technology. European Commission, Brussels vol. 21 (1), pp. 25–44 (2001)

Chapter 14

Brain-Computer Interface (BCI): Types, Processing Perspectives and Applications

Konstantinos J. Panoulas, Leontios J. Hadjileontiadis, and Stavros M. Panas

Department of Electrical & Computer Engineering
Aristotle University of Thessaloniki
GR 54124 Thessaloniki, Greece
Tel.: +302310 996340; Fax: +302310 996312
{panoulas,leontios,spanas}@auth.gr

1 Introduction

A Brain-Computer Interface (BCI) is a system that aims to create a direct communication channel between the brain and a computer, with the purpose of transmitting messages and commands. Such systems utilize well defined underlying correlations between certain mental activities and electrophysiological signals associated with the brain. Depending on the positioning of the sensors used to record the aforementioned signals, BCI systems can be categorized as noninvasive when sensors are placed on the scalp, measuring either the electrical potentials produced by the brain which is called electroencephalography (EEG) or the magnetic fields with a technique called Magnetoencephalography (MEG); semi-invasive when electrodes are placed on the exposed surface of the brain in a practice called electrocorticography (ECoG); and invasive, when micro-electrode arrays are placed directly into the cortex.

Noninvasive systems, which currently lie in the research focus, mainly utilize EEG recordings because they are easily acquired by a plethora of commercial off-the-shelf devices at a relative low price. Every technique that has been based on EEG recordings has to address challenges that are inherent to EEG, specifically its poor spatial resolution, which results in interference from unwanted signals, and its low signal to noise ratio. MEG recordings have been utilized to drive BCI systems, taking advantage of MEG higher spatiotemporal resolution, but they place a need for sensitive sensors and magnetically shielded rooms and usually have limited their research potential.

Semi-invasive systems using ECoG provide better spatial resolution and signal-to-noise ratios than EEG at the cost of an invasive procedure called craniotomy. The biophysical characteristics of EEG and ECoG recordings are similar in many respects and systems based on ECoG exploit the same underlying neurophysiologic mechanisms as EEG based systems, which lead to a common approach from a signal processing point of view.

Invasive systems use implantable micro-electrodes placed into the cortex in a highly invasive procedure which has led research to focus on animals, mainly monkeys and rats, even though such systems have been demonstrated in humans. While recordings acquired from such electrodes have very high spatial resolution and signal-to-noise

G.A. Tsihrintzis and L.C. Jain (Eds.): Multimedia Services in Intel. Environ., SIST 3, pp. 299–321.
springerlink.com

ratio, which results in significantly improved system performance, degradation of the signal quality over time due to brain tissue response to the electrodes presents an important challenge. Micro-electrode recordings exhibit large differences from EEG/ECoG recordings and require different signal processing approaches.

2 Principles of Noninvasive BCI operation

For noninvasive BCI systems the most common electrophysiological signal is the EEG which in order to be used as a basis for a BCI system, must be correlated with a mental process, that is either conscious or can be affected consciously so that it can be used to represent intention. Furthermore, it must be able to be well characterized for an individual, so that it can be reliably tracked and detected. There are several signals and corresponding mental processes that fulfill the aforementioned requirements that create corresponding strategies for the creation of BCI systems. These are described in the subsequent sections.

2.1 Event Related Potentials

For EEG recordings, some sensory stimuli and cognitive processes trigger stereotyped brain responses that are called Event Related Potentials (ERPs) [1]. ERPs which are recorded following an external stimulus are called Evoked Potentials (EPs) [2], [3], whereas when sensory organs are stimulated, they are called Sensory Evoked Potentials (SEPs) with the most important SEPs being the Visual Evoked Potentials (VEPs) and the Auditory Evoked Potentials (AEPs). The ERP waveforms are usually described by their amplitude and latency, thus ERPs are categorized as positive or negative (represented by the letters P or N respectively) and identified by the number of milliseconds after the trigger event that they occur. For example, a well studied ERP is the P300 [4], which is a positive deflection in voltage which can be witnessed roughly around $300\ ms$ after the triggering event. An example of an ERP recorded from channel C4, according to international 10/20 standard, which contains several components along with a dominant P300, is shown in Fig. 1.

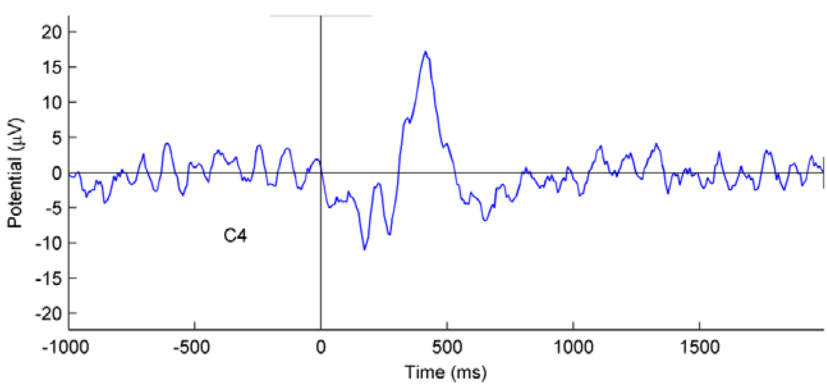


Fig. 1. An ERP containing a P300 component recorded by an electrode at the C4 position according to the international 10/20 standard. The actual component is the positive deflection that starts at about 300ms and ends at about 550ms. The time is zero at the presentation of the stimulus.

The P300 is a composite wave [5] which is recorded after stimuli that requires information processing, and one of its components occurs when the subject detects an occasional "anticipated" stimulus from a set of regular stimuli, a framework which is usually referred to as the "oddball paradigm". It has been documented [6] that the amplitude of this component is increased when the stimulus is less anticipated, but still remains relatively small and usually requires averaging of multiple recordings. The P300 waves of VEPs were used for the first time in a BCI in 1988 by Farwell and Donchin [6] in an application that displayed a 6 by 6 matrix of characters to the user where various rows or columns were highlighted. When a row or column that contained the character the user has selected was highlighted, a P300 response was elicited, since this was the target character, i.e., the "oddball paradigm". The matrix used in this first P300 BCI system is shown in Fig. 2. A talk command was contained because the system was connected with a speech synthesizer that had the ability to sound the word that had been selected.

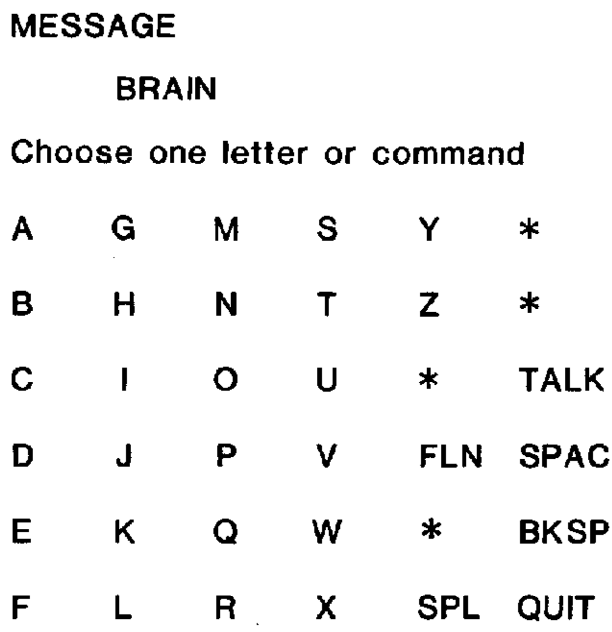


Fig. 2. The matrix that was used for the first BCI application using visually evoked P300 ERP [6].

2.2 Steady State Visual Evoked Potentials

A distinct case of VEPs stems from visual stimulus that is modulated to a specific frequency ranging from 3.5 to75 Hz, with most often used the 13 Hz. The brain response, which is called Steady State Visual Evoked Potential (SSVEP), has the same fundamental frequency as the stimulating frequency and usually includes harmonics. An example of the spectral content of an EEG recording during the presentation of a visual stimuli modulating at 7Hz is depicted at Fig. 3.

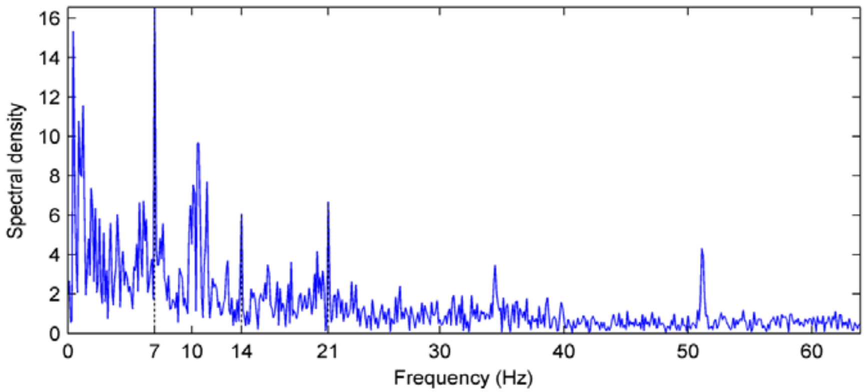


Fig. 3. The spectral content of an EEG recording during the presentation of a visual stimuli modulating at 7Hz. The fundamental frequency of 7 Hz is visible so are the first and second harmonics at 14 and 21Hz, respectively.

When two or more frequencies are simultaneously modulating the visual stimuli the frequency that has the focus of the user's gaze is prevalent and thus SSVEPs can be used as a basis for a BCI design, as successfully was demonstrated in 2000 by Middendorf *et al.* [7]. This BCI application had two virtual buttons on a computer screen, flashing at different frequencies and used SSVEPs to allow users to select the button they desired by looking at it.

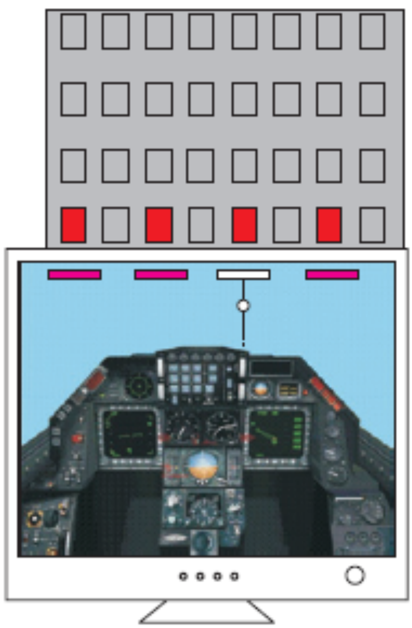


Fig. 4. An SSVEP BCI used to control a computer game [8]. The flashing LED array was positioned above the computer screen.

One limitation of the BCI systems based on SSVEPs is that users must have good voluntary control of their eye movements but overall it has been suggested that such systems are more feasible than others [8]. An example of a visual stimulus interface which has 32 buttons, four of which are used to control a computer game which involves flying an airplane, is depicted at Fig. 4.

2.3 Event-Related De/Synchronization

Internally or externally paced events are linked not only with ERPs but also with an ongoing change in the EEG. It has been shown that certain events or mental processes can reduce or desynchronize ongoing alpha wave brain activity [9]. These event-related phenomena result in specific changes in the ongoing EEG activity and consist, in general terms, of power increases or decreases in certain frequency bands [10]. A decrease in the spectral power is called event-related desynchronization (ERD), while an increase is called event related synchronization (ERS). These phenomena are considered to be due to a decrease or an increase in the synchronization of the underlying neuronal populations, respectively. ERD and ERS events for a frequency band of interest are measured as a function of the power of that frequency band in the period after the event (denoted as A) and the power of the same frequency band in a reference period (denoted as R), by the following formula as defined in [10]:

$$ERD/ERS(\%) = \frac{A-R}{R} * 100 \tag{1}$$

An ERD of the upper alpha and lower beta frequency bands $(8 - 12\ Hz)$ and of the beta band $(20 - 24Hz)$, localized close to the corresponding sensorimotor cortex areas[1] has been linked with ongoing voluntary movements [10], [11]. The same ERD has been witnessed in imaginary movements [12], [13] a phenomenon which has been utilized for the creation of the Graz brain-computer interface II, the first BCI application that uses imaginary hand and foot movements to distinguish between three different EEG patterns: planning or preparation of movement of the left index finger, right index finger and right foot.

A common setup for a simple two-class BCI is the recording of three channels (C3, Cz and C4 according to the international 10/20 standard, depicted in Fig. 5. Channel C3 is located above the left sensorimotor cortex; channel C4 above the right sensorimotor cortex and CZ lies between the two.

The ERD from an imaginary left hand movement is depicted in Figs. 6 and 7. There, the spectrograms of channels C3 and C4 are depicted, respectively, recorded during an experiment where the subject was instructed at $t = 3\ s$ to imagine a left hand movement. Since the left hand is controlled by the right sensorimotor cortex, channel C4 demonstrates ERD (suppression of the mu frequencies $(9 - 12)\ Hz$), while channel C3 remains unaffected. Similarly, Figs. 8 and 9 illustrates the spectrograms of channels C3 and C4 during the same experiment but when the subject was instructed to imagine a right hand movement. Channel C3 exhibits ERD while channel C4 remains unaffected.

---

[1] Movement of the left side of the body is controlled by the sensorimotor cortex located in the right side of the head and vice-versa.

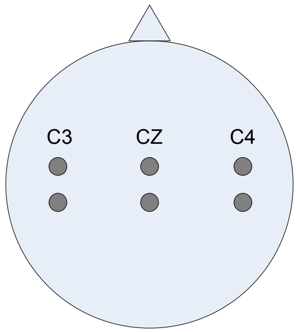


Fig. 5. A common electrode positioning scheme for the implementation of a simple two class BCI based on imaginary left and right hand movements. The locations of the channels are identified (left to right) as C3, CZ and C4 according to the international 10/20 standard.

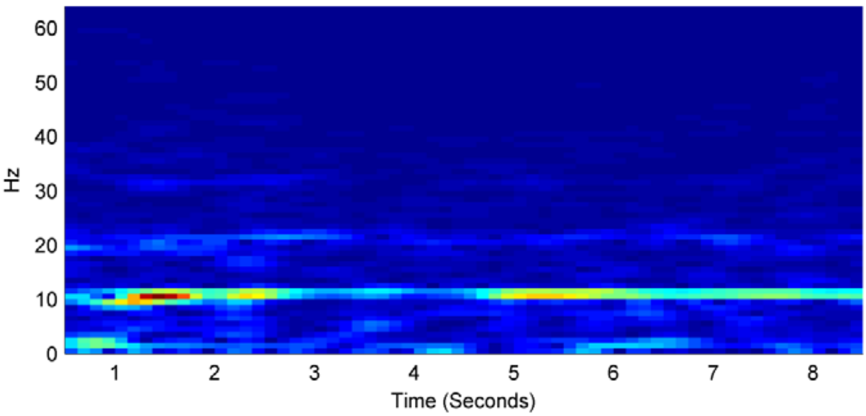


Fig. 6. The spectrogram from channel C3 during an imagined left hand movement that starts at about $\boldsymbol{t = 3\ s}$; as expected, the mu rhythm remains unaffected.

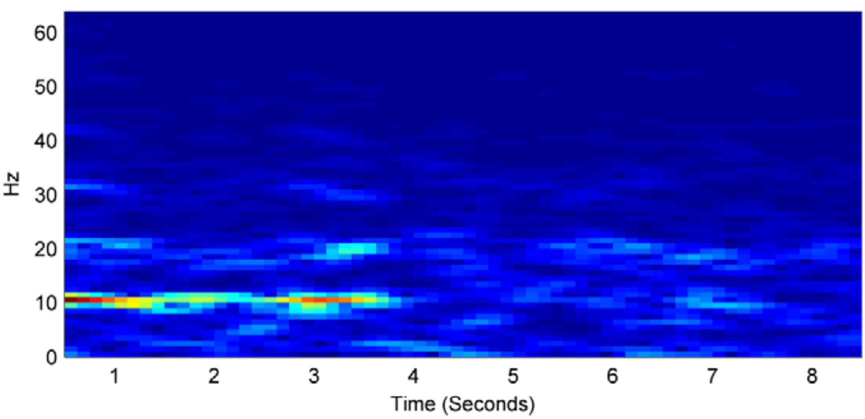


Fig. 7. The spectrogram from channel C4 during an imagined left hand movement that starts at about $\boldsymbol{t = 3\ s}$; as expected, the mu rhythm is suppressed while the imaginary movement lasts.

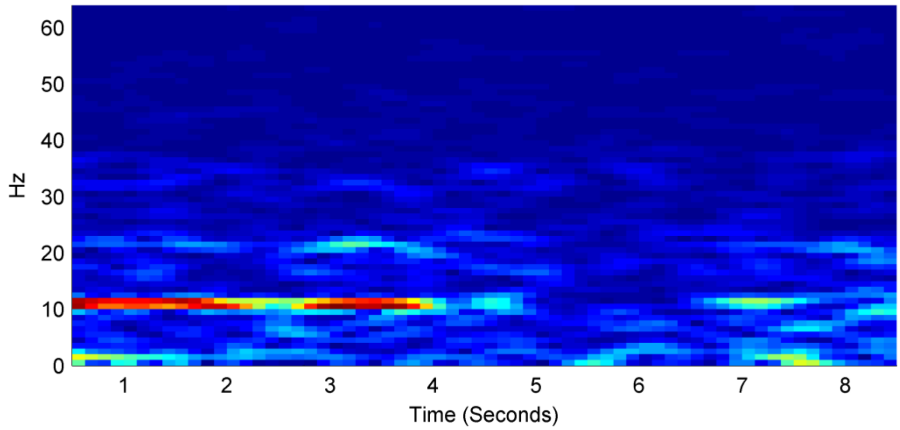


Fig. 8. The spectrogram from channel C3 during an imagined right hand movement that starts at about $\boldsymbol{t = 3\ s}$; as expected, the mu rhythm is suppressed while the imaginary movement lasts.

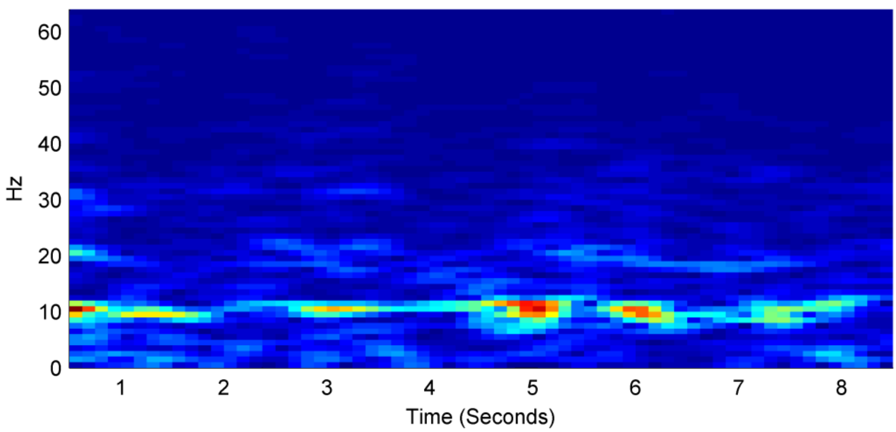


Fig. 9. The spectrogram from channel C4 during an imagined right hand movement that starts at about $\boldsymbol{t = 3\ s}$; as expected, the mu rhythm remains unaffected

2.4 Slow Cortical Potentials

Slow Cortical Potentials (SCP), which are also known as DC potentials, are surface recorded waves with frequencies less than $2\ Hz$ that are linked with various cognitive events, for example anticipation, cognitive preparation and motivational states of apprehension and fear; thus, individuals can be trained to modify SCPs by using feedback. They manifest as positive or negative shifts and physiologically are presumed to reflect the extent to which apical dendrites of the cortical pyramidal cells are depolarized. In an effort to identify the cortical sources of the SCP phenomena fMRI imaging was used [14] to locate the areas of the brain that are activated during self-regulation of SCPs. Generation of negativity was accompanied by a widespread activation in central, pre-frontal, and parietal brain regions, as well as the basal ganglia, while generation of positivity was accompanied by widespread deactivation in several cortical sites, as well as activation primarily in frontal and parietal structures, as well as insula and putamen. This means that negative shifts are associated with increased cortical activation and positive shifts are associated with decreased cortical activation. The activation and deactivation of cortical areas during negativity and positivity is depicted at Fig. 10, where areas painted red are indicative of significant activation and areas painted green are indicative of deactivation when compared to a baseline.

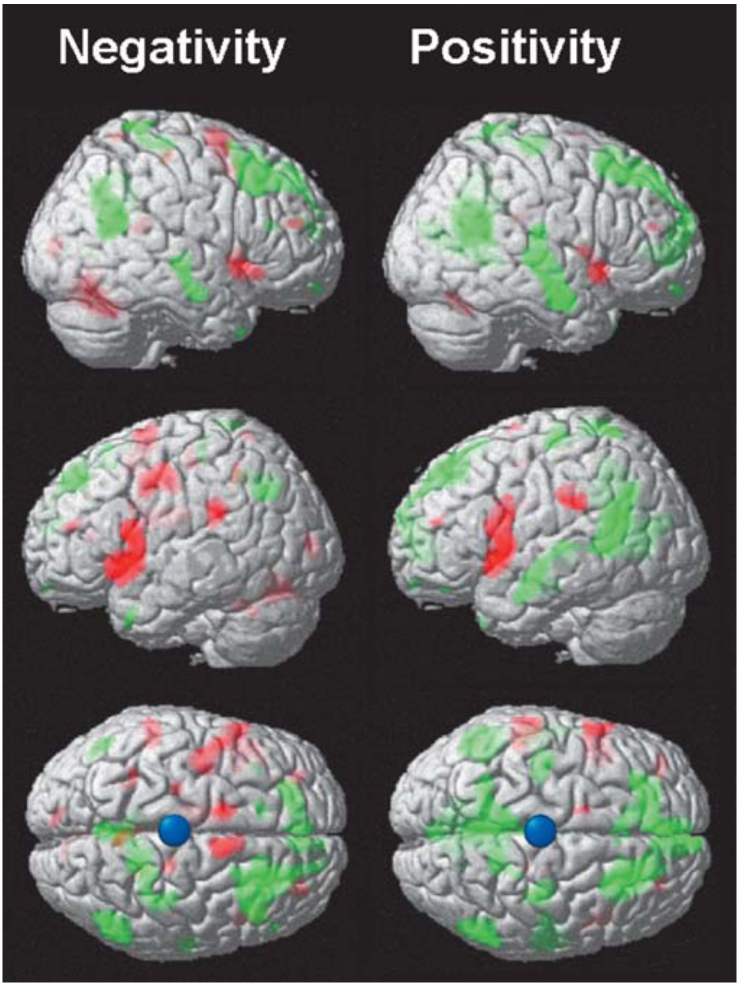


Fig. 10. Activation and deactivation of cortical areas during negativity and positivity. Areas painted red are indicative of significant activation and areas painted green are indicative of deactivation when compared to a baseline. The positivity task shows deactivations in frontal and temporo-parietal areas. The blue dot in the bottom pictures marks the active central electrode position used for feedback. Image was adapted from [14].

An example of EEG recordings during self regulation of SCPs is shown in Fig. 11. The recordings were acquired from electrode location C4 according to international 10/20 standard where the increased activation and deactivation are evident.

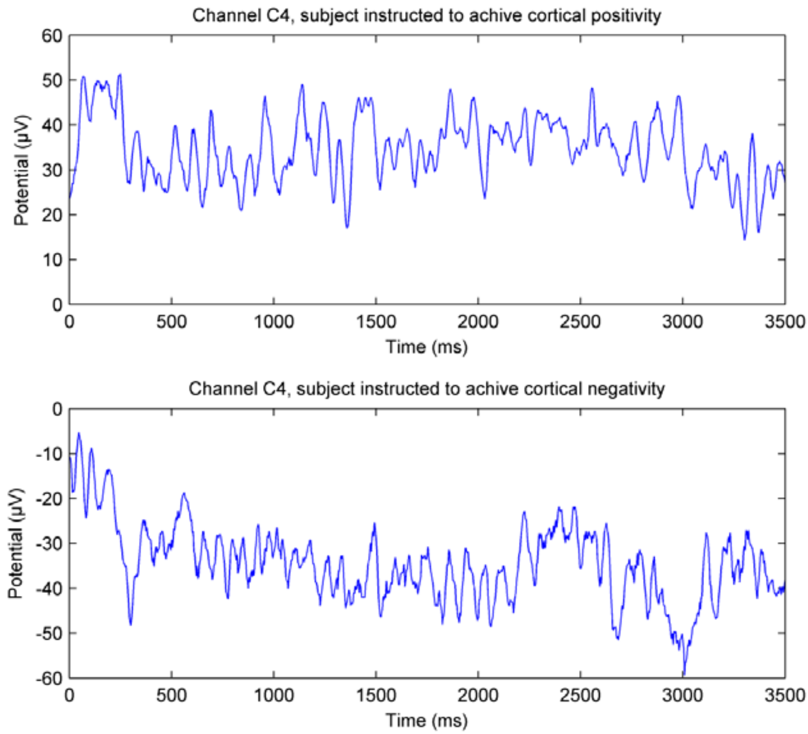


Fig. 11. Example of EEG recordings during self regulation of SCPs. The recordings were from electrode location C4 according to international 10/20 standard where the increased activation and deactivation are evident.

3 State-of-the-Art in EEG-Based BCI Signal Processing

All EEG-based BCI classes have to face the problem of separating the control signal from interfering noise signals which have two sources: non-EEG artifacts, such as recording noise, power line interference, eye movement, eye blinking, EMG and ECG; and EEG signals that are not used as control signals. The two noise classes differ from the control signal either in their frequency distribution, their topographical location of their source in the brain or both. For example, eye movement signals have maximal frequency content in low frequencies ($< 5\ Hz$) and are located over anterior head regions. Similarly, the visual alpha rhythm lies inside the frequency range of most BCI control signals and it is more evident in the parieto-occipital cortex [15]. While the elimination of noise sources whose frequency distributions lie outside the frequency range of the control signal can be easily implemented by filtering, the elimination of the other noise sources poses a far more difficult problem and necessitates the use of advanced signal processing methodologies. After the control signal is separated from unwanted noise, it is fed to a classifier to convert it to a command signal, which can be used for control and to a user's feedback visualization component, which helps the user refine the control signal. A block-diagram of a typical EEG-based BCI application is illustrated in Fig. 12. A description of the most commonly used methods for signal separation and classification follows.

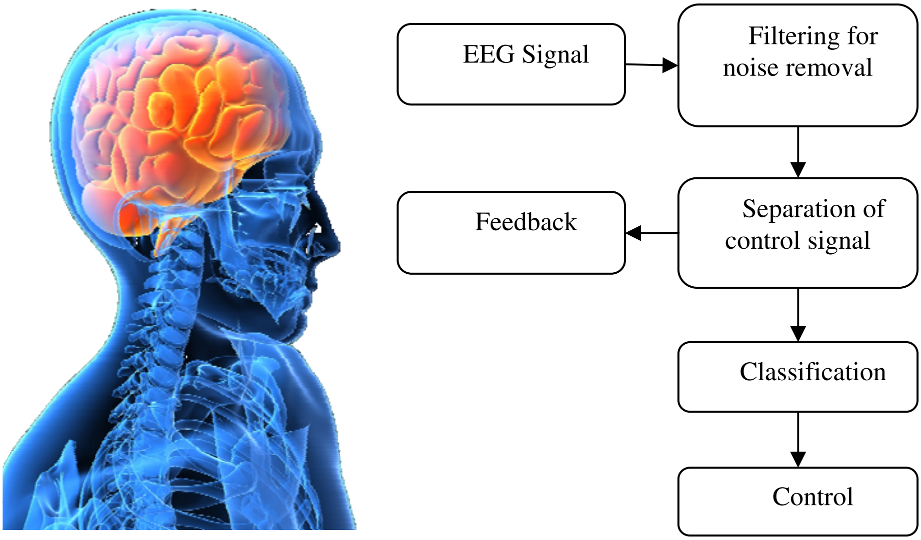


Fig. 12. A block-diagram of a typical EEG-based BCI application. After the EEG signal is acquired, it is filtered for basic noise removal. The control signal is then separated from artifacts and contamination by EEG signals unrelated to the target mental process, and a feedback based on the raw control signal is presented to the user. The control signal is also entered to a classifier that generates the control commands which are also presented.

3.1 Laplacian Spatial Filtering

The Laplacian method emphasizes electrical activity which originates by radial sources immediately below the electrode by calculating the second derivative of the instantaneous spatial voltage distribution. High spatial resolutions can be achieved with this method by using many electrodes. To calculate the Laplacian derivations, a finite difference method is used which approximates the second derivative by subtracting the mean activity from surrounding electrodes of the electrode of interest using the following formula defined in [15]:

$$\boldsymbol{V}_i^{LAP} = \boldsymbol{V}_i^{ER} - \sum_{j \in S_i} g_{ij} \boldsymbol{V}_j^{ER}, \tag{2}$$

where

$$g_{ij} = \frac{\frac{1}{d_{ij}}}{\sum_{j \in S_i} \frac{1}{d_{ij}}}, \tag{3}$$

S_i is the set of electrodes which surround the i^{th} electrode and d_{ij} is the distance between electrodes i and j.

The effects of the Laplacian filtering vary with the distance between the electrodes thus a uniform Laplacian filter may not be appropriate in all BCI cases. Even though the method was one of the first that was used, recent studies [16] have shown that it performs comparably to more recent methods as Independent Component Analysis (see §3.3).

3.2 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a linear transformation that expands the data into a set of orthogonal components ordered by their variance. The fact that the transform kernels are orthogonal achieves maximum decorrelation and can separate the signal and noise components more efficiently than transforms which use kernels that are independent of the data. In case of a dataset $\boldsymbol{X}$ which is a $m \times n$ matrix, where m is the number of measurements and n is the number of samples for each measurement, the analysis of $\boldsymbol{X}$ of into its principal components $\boldsymbol{P}$ is done using the formula:

$$\boldsymbol{Y} = \boldsymbol{P} \cdot \boldsymbol{X} \tag{4}$$

In order to calculate $\boldsymbol{P}$ matrix, the constraint that the covariance matrix of $\boldsymbol{Y}$, $\boldsymbol{C_Y}$ defined as $\boldsymbol{C_Y} = \boldsymbol{Y} \cdot \boldsymbol{Y^T}$, is diagonal is applied. It is deduced that the principal components can be calculated by the eigenvectors of $\boldsymbol{C_X}$ which is the covariance matrix of the measurement data matrix $\boldsymbol{X}$. PCA decomposition has been used together with wavelet analysis and has been suggested that it can be used to extract meaningful components for analysis with some limitations [17].

3.3 Independent Component Analysis (ICA)

The concept of Independent Component Analysis (ICA) lies into decomposing a multivariate signal into additive source signals with the assumption that they are statistically independent. Considering a multichannel signal as $\boldsymbol{x}(n)$, and the signal components as $x_i(n)$, the $x_i(n)$ signals are independent if:

$$p_X\big(\boldsymbol{x}(n)\big) = \prod_{i=1}^{m} p_x(x_i(n)), \forall n \tag{5}$$

where $p_{\boldsymbol{X}}$ is the joint probability distribution, $p_x(x_i(n))$ are the marginal distributions and m is the number of the signal components.

The ICA is used to solve the general problem of blind source separation (BSS) which is to estimate and recover the source signals using only the information of their observed mixtures; a problem which takes the name of its acoustic analog thus referred to as the "cocktail party problem". The separation is made by using a separating matrix $n \times m$ where n is the number of recordings, hence in EEG-based BCI problems the number of the electrodes, and m is the number of independent sources, with $n \geq m$.

Two main strategies have been used to quantify the statistical independence between the acquired EEG signals. The first is by using Mutual Information (MI) of $\boldsymbol{x}$ which is zero if and only if the components of $\boldsymbol{x}$ are mutually independent and is strictly positive otherwise [18]. Two ICA algorithms (widely used in BCI applications), which use this strategy are FastICA [19] and INFOMAX [20]. The second is through the use of higher-order cumulants, since it is known that if at least two components of $\boldsymbol{x}$ are statistically independent then all cumulants involving these components are zero.

The use of ICA for the removal of a wide variety of artifacts from EEG recording in non-BCI applications was demonstrated first [21] followed by a widespread use in all BCI classes. In P300 based BCIs, ICA was used both to remove unwanted artifacts [22] and to separate target and non-target ERPs [23] by the selection of meaningful independent components using *a priori* physiological knowledge. In addition, ICA was used for SSVEP-based BCIs, where the more conventional methodology of Fourier analysis is mainly used to aid the selection of the signal and reference channels, as proposed by Wang *et al.* in [24]. Applications based on imaginary movements and the detection of ERD are the best candidates for the application of ICA, since the two sensorimotor cortices lie on different sides of the head and are thought to be statistically independent. ICA has been used as a spatio-temporal filter [25] in a BCI paradigm which uses 59 electrodes, and as a means to separate the two sources in a much simpler BCI case with only three channels.

3.4 Common Spatial Patterns (CSP)

The common spatial patterns method is primary used in BCI systems based on ERD/ERS phenomena and since it is parallel by nature it is well fitted for on-line data processing. The method was introduced in the field of EEG analysis by Koles *et al.* in 1990 [26] and was used with respect to EEG-based BCI applications by Muller-Gerking *et al.* in 1999 [27]. The method works by constructing very few new time-series whose variances contain the most discriminative information, that are subsequently used to feed a classifier.

Let $\boldsymbol{V}_a^i$ denote the raw data of trial i, under the condition a which is represented as a $N \times T$ matrix where N is the number of channels and T is the number of samples. Let R_a^i represent the normalized covariance matrix of $\boldsymbol{V}_a^i$ calculated by

$$R_a^i = \frac{V_a^i \cdot \overline{V_a^i}}{trace\left(V_a^i \cdot \overline{V_a^i}\right)}, \tag{6}$$

and let R_b^i represent the normalized covariance matrix of $\boldsymbol{V}_b^i$. The normalization is done to eliminate inter-trial variations in the absolute values of the standard deviation. Next, the normalized covariances are averaged over trials thus creating matrices R_a and R_b, respectively. Afterwards, the composite covariance matrix is created $R_c = R_a + R_b$ and its eigenvectors matrix B_c and eigenvalues matrix λ are computed. The whitening transformation matrix

$$W = \lambda^{-\frac{1}{2}}\overline{B_c} \tag{7}$$

transforms R_a and R_b to

$$S_a = WR_a\overline{W} \tag{8}$$

and

$$S_b = WR_b\overline{W}. \tag{9}$$

Since S_a and S_b share the same eigenvectors, the eigendecomposition of S_a or S_b gives the orthonormal matrix U. The projection of whitened EEG epochs on U gives feature vectors that are optimal in the least squares sense for discriminating between the two populations and can be seen as time-invariant EEG source distribution vectors.

The method has been used in [28] with high classification results, and in a comparison with the most prominent algorithms for spatial filtering [16] it yielded better results compared to all other algorithms. Furthermore, the utilization of CSP in [29] demonstrates the application of the method when more than two classes (conditions) exist.

3.5 Linear Discriminant Analysis Classifier

One of the oldest but widely used classification processes is the Linear Discriminant Analysis (LDA) [30]. LDA classifiers are commonly used in SSVEP-based BCI paradigms [31], but due to their simplicity and high performance are used in all BCI classes [32].

In the LDA process, a classification criterion (Bayes' rule) is used to minimize the total error of classification (TEC) tending to make the proportion of object that it misclassifies as small as possible. In other words, TEC should be thought of as the probability that the rule under consideration will misclassify an object. If there are g groups, the Bayes' rule is to assign the object to group i where $P(i|\mathbf{x}) > P(j|\mathbf{x}), \forall j \neq i.$, where $\mathbf{x}$ denotes a set of features from measurements (feature vector). With the help of Bayes' theorem and if we assume that each group has multivariate normal distribution and all groups have the same covariance matrix ($\boldsymbol{C}$), we get what is called as LDA formula,

$$f_i = \boldsymbol{\mu}_\mathbf{i}\mathbf{C}^{-1}\mathbf{x}_\mathbf{k}^\mathbf{T} - \frac{1}{2}\boldsymbol{\mu}_\mathbf{i}\mathbf{C}^{-1}\boldsymbol{\mu}_\mathbf{i}^\mathbf{T} + \ln(p_i), \tag{10}$$

where $\mathbf{x_k}$ represents the features of object k; p_i is the prior probability about the group i known without making any measurement (if we do not know the prior probability, we just assume it is equal to the total samples of each group divided by the total samples); $\boldsymbol{\mu}_\mathbf{i}$ is the mean of features in group i, which is average of $\mathbf{x_i}$. Using (10), LDA assigns object k to group i that has maximum f_i.

3.6 Support Vector Machines

Support Vector Machines (SVM) is a category of classification methods which use supervised learning to separate two different classes of data. The idea behind SVM is to construct a hyperplane, described by a weight vector $\boldsymbol{w}$ and a bias value b, which will separate the two different classes of data, using a training set of l samples with data vectors $\boldsymbol{x}_i$ and corresponding class labels y_i, where $\boldsymbol{x}_{i\in}R^N$ and $y_i \in \{-1,1\}$ [33]. In the test phase, the class of a new data vector $\boldsymbol{y}$ can be predicted by projecting $\boldsymbol{y}$ on the weight vector $\boldsymbol{w}$ as follows:

$$f(\boldsymbol{y}) = \boldsymbol{w} \cdot \boldsymbol{x} + b \tag{11}$$

The sign of this projection will identify the predicted class label. Because there are several possibilities for the selection of hyperplanes separating the data into two subsets, a suitable criterion for the selection of a weight vector is the maximization of the separation margin γ between the two classes. To describe this optimal hyperplane, the vectors which lie on the margin (called support vectors) are only necessary.

4 Invasive BCI Systems

As it was documented through experiments initially on monkeys and subsequently on humans, electrical activity generated by individual cortical or subcortical neurons that are associated with movements can be used as a basis for the creation of an invasive BCI system. Such neurons have been located both in the primary motor cortex and in the posterior parietal cortex. Since invasive BCIs tap directly into the brain's motor commands processing functions, they do not require extensive training to control the output, even though performance increases are witnessed. Different approaches have been suggested to record this neuronal activity. Some approaches use local field potentials, which are recordings of the summation of electrical activities of neurons which reside inside a particular volume of tissue, while others use recordings from small or large samples of individual neurons from single or multiple units.

Even though some studies have claimed that recordings from small numbers of highly tuned neuronal groups have been sufficient for good BCI performance, the fact that the surgical procedure of implanting a microelectrode results in a partially random selection of neurons, means that the existence of highly tuned neurons in a typical recording is rare. Therefore, the recordings of large samples of neurons are preferable either to increase the probability of detecting highly tuned neurons or to increase the accuracy and reliability of such systems by reducing individual neuron firing variability.

The acquisition of recordings either from individual neurons or from local field potentials presents major technological challenges which are a result of the broad issue of biological compatibility. Current microelectrode designs typically enable recordings for a duration of months but quality often deteriorates possibly due to electrode encapsulation by fibrous tissue and death of the cells that are being recorded. Higher recording durations have been reported for certain examples and certain species, but high quality, long term recordings still remain elusive. Apart from biocompatibility issues, research into microelectrode designs focuses on the creation of electrode matrices capable of simultaneously recoding hundreds of neuronal signals, to satisfy the need for large sample recordings, like the Utah electrode array probe [34], which is depicted in Fig. 13.

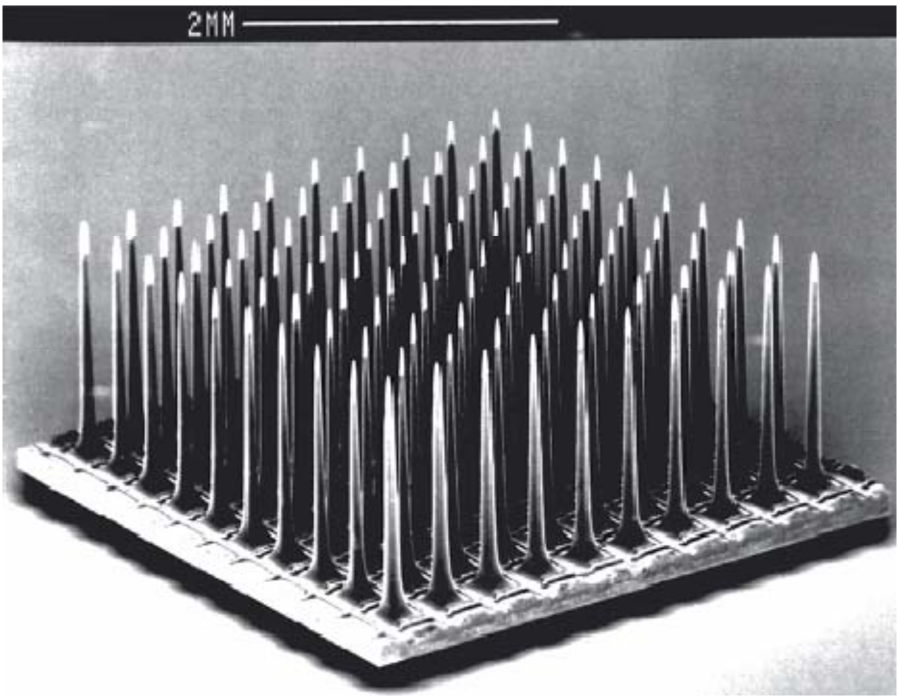


Fig. 13. The Utah electrode array consisting of 100 electrodes that extend 15*mm* from the $10 \times 10\ mm$ substrate. With permission from *Encyclopedia of Sensors* [34].

5 Resources for BCIs

Research and applications in BCIs and EEG signal processing has been greatly aided by the availability of two high quality tools, namely the EEGLab and BCI2000. EEGLab is an active open source project that is currently being developed by the Swartz Center for Computational Neuroscience (SCCN) of the Institute for Neural Computation at the University of California San Diego (UCSD) in collaboration with the

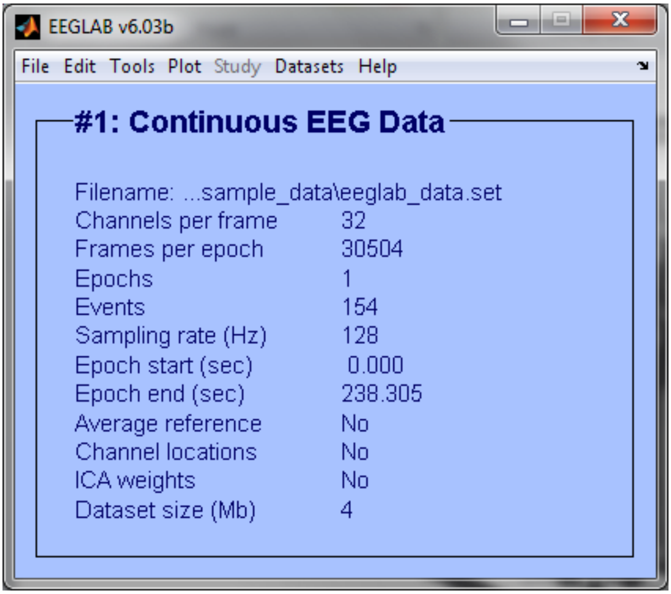


Fig. 14. The main interface of EEGLab

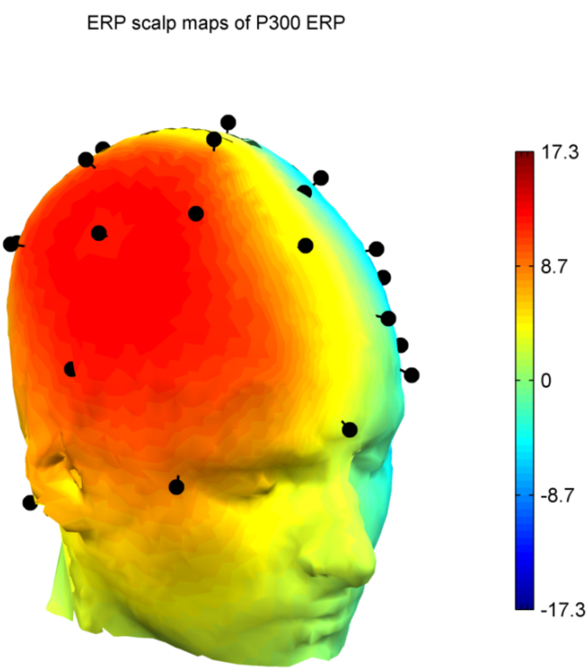


Fig. 15. A 3D scalp map of an ERP signal with a P300 component within the EEGLab

CNRS CERCO laboratory in France. It is mainly focused on processing continuous and event-related EEG, MEG and other electrophysiological data, thus it is very well suited for analysis of signals related to BCIs. Many of the signal processing tools commonly used for BCIs and some useful visualization schemes are implemented natively or by plug-ins for EEGLab. The main interface of EEGLab and a 3D scalp map of an ERP signal with a P300 component are depicted in Figs. 14 and 15.

BCI2000 is developed by the BCI R&D Program at the Wadsworth Center of the New York State Department of Health in Albany, New York, USA, in collaboration with the University of Tübingen in Germany. BCI2000 is a complete open source BCI research system that consists of several modules and can operate either as based on the P300 evoked potential or based on the ERD of the mu rhythm.

6 Applications of BCI Research

The main scope of BCI research lies with medical applications directed to individuals with inabilities that require an alternative communications or control channel. Patients with amyotrophic lateral sclerosis (ALS) or locked-in syndrome have been able to use BCI systems with success to control aspects of the environment or communicate. BCI applications have also been used to control robotic limbs by patients with amputations. Furthermore patients with spinal cord lesions have used BCIs in conjunction with functional electrical stimulation (FES) devices.

6.1 Invasive BCIs

Commercial applications oriented towards medical uses of BCIs have been created with non-invasive, as well as with invasive BCI systems. In the area of invasive BCI applications a company named BrainGate is currently engaged in clinical trials for an implantable system that decodes imagined limb movements to control prosthetic arms, wheelchairs or personal computers [35].

6.2 Noninvasive BCIs

Contrary to invasive systems, non-invasive BCI system are currently reliable enough to be used as an alternative means of communication for patients outside dedicated laboratories, at the ease of their homes [36]-[43]. Although this requires the training of caretakers in the application of electrodes and the recognition of recording problems, such advances are already a reality. The creation of systems which use reduced montages (eight channels or less) and the use of "dry electrodes" that do not require gel will helps address these concerns.

Commercial applications have already been created to demonstrate that BCI technology can reliably allow patients to affectively control movement and day-to-day functions in a smart home by using an evolved P300 speller. A virtual 3D representation of a virtual home along with control masks have been defined to allow realization of everyday tasks, such as moving around, opening doors, watching television and listening to music. The system was created by a company named g.Tec (www.gtec.at) and it is depicted in Figs. 16-18.

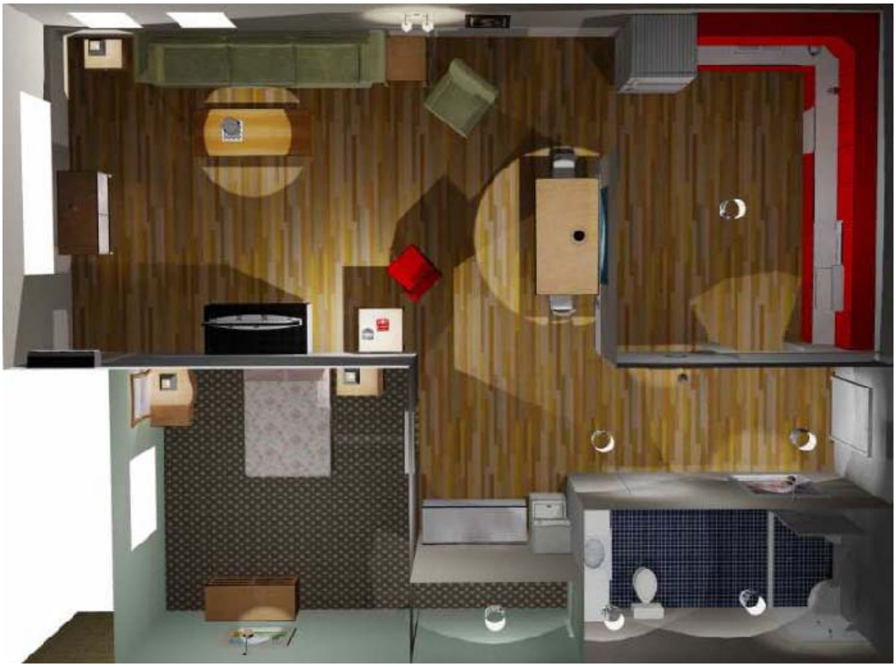


Fig. 16. The layout of the smart virtual home (with permission from g.Tec)

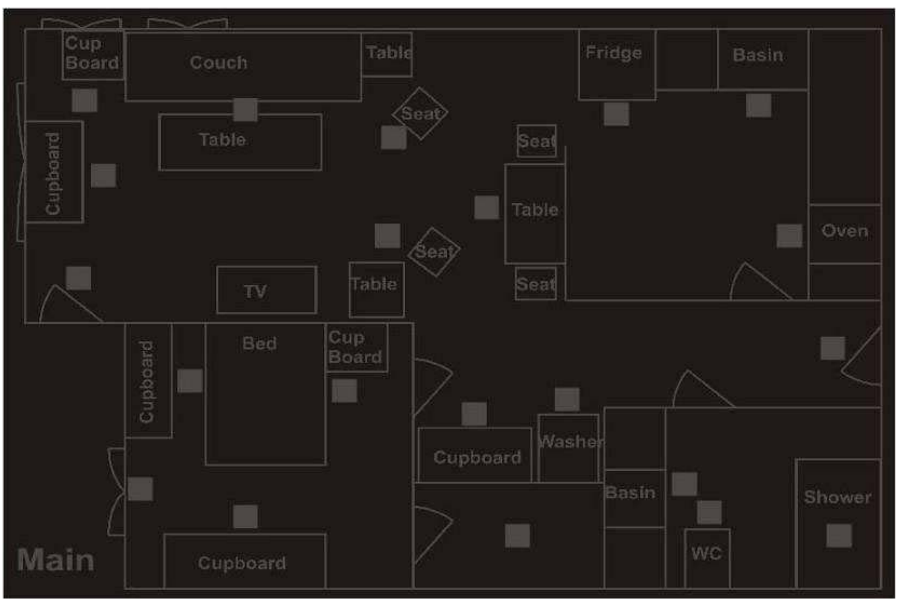


Fig. 17. The location selection matrix for the virtual home depicted in Fig. 16. The user of the BCI system can select the object to interact with by focusing on it; the system recognizes the selection using the principles behind the P300 BCIs (with permission from g.Tec).

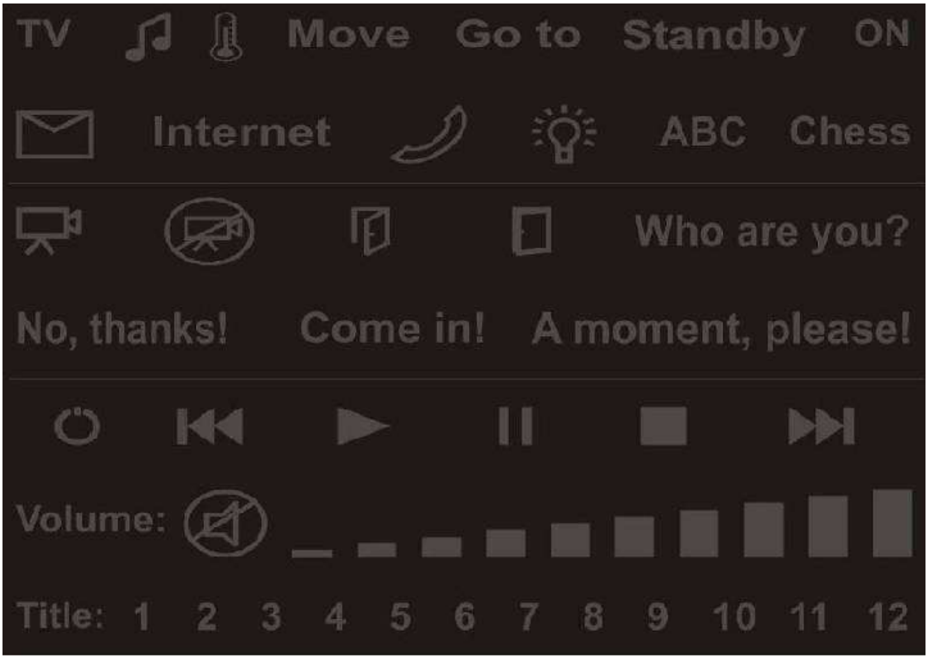


Fig. 18. The selection matrix for the television (with permission from g.Tec)

7 Concluding Implications

There is huge potential for the BCI applications in the form of products in so many different areas. The most beneficial application for the humans will be to help the disabled people. Paralyzed people will be able to move more easily, they will be able to perform most of the normal day activities with limited help/supervision from another individual. Another area where BCI can be very effectively used is in the field of surgery. Recently, there have been surgeries performed where the surgeon is one part of the world and the patient is in another part of the world and using robotic arms the procedure was performed. But in these kind of surgeries, the surgeon has to control the robotic arm using some controls at his end. If BCI technology can be effectively used at the surgeon's end that will be huge breakthrough in the field of surgery. The surgeon will use some kind of non-invasive BCI device and just looking at the patient's live images on computer screen and by the thoughts that go through his mind, he would be able to control the robotic arms to perform the procedure. BCI devices like the ones which have been developed by Emotiv (www.emotiv.com), the video gamers can play the games on computer using their thoughts and simple body movements, which will be step ahead from Wii. BCI technology has not been explored to much extent yet for law enforcement agencies but there is lot of applications which can be developed. MEG and MRI can construct the images of a person from human brain and display on the computer. So during interrogation the sketches from the arrested person's brain can be developed which can lead to further investigation about other people involved.

There are many technical challenges currently faced by this technology. One of the technical challenges is of calibration. The non-invasive device needs to be calibrated before every use and this calibration step can take some time. So developing a fast calibrating system is a challenge. Another challenge is use of use. Some BCI devices require some training before they can be used effectively especially in case of disabled people and there is large range of complexity and different brain state from one individual to another. So developing a BCI which can be used to a vast population of people without extensive training is a challenge. Another big challenge facing brain-computer interface researchers today is the basic mechanics of the interface itself. Non-invasive BCI blocks some of the signals from the brain and it has been a challenge for the developers. Another challenge is that these devices cannot capture all the signals from the brain. As brain uses electrochemical signals, these devices can only sense the electric signals and the chemical processes cannot be read by these devices, so the signal received might not be a complete one. In addition, selection of the dependent variable (i.e., external stimulus or predicted state) poses an interesting issue in modeling and classification processes within a BCI system. Finally, there are many legal and social challenges of BCI devices. What information can be shared and what information the BCI device extracts. There will be privacy issues associated with the BCI devices. Some of the challenges will be that what information that device collects, how it is used, how it is saved, where it is saved and who will have access to it.

Additional issues that acquire attention are [44]: the nonstationarity of the neurophysiological changes in time and space; how to cope with adaptability in the subject's and the BCI controller's site; BCI generalization to a variety of tasks; how to optimally bridge the time scale of spike events (msec) with the time scale of behaviors (sec); enhancement of BCI robust behavior through more sophisticated self-organizing adaptive principles.

In any case, BCI systems could serve as experimental platforms were brain theories could be tested, revealing all the ingredients of cognitive experimentation with high resolution, providing synchronous measurement access both at the input (e.g., spike trains) and at the behavior level. Collaboration amongst engineers, neuroscientists, physical scientists, and social and behavioral scientists, towards the integration and convergence of engineering tools and methods in the areas of sensors and signal processing, noninvasive and minimally invasive recording techniques from the brain and the peripheral nervous system, neural tissue engineering, neural imaging, nonlinear dynamics, chemical and biological transport, computational neuroscience and multiscale modeling, nano/micro technological neuroscience, control theory, systems integration, and robotics share the same endeavor; that is, to permit control of movement where normal neural pathways do not exist. Transformational solutions being pursued are leading to better understanding of the central and peripheral nervous systems and pushing forward the frontier of scientific discovery.

As a bottom line, the evolution of the BCI systems proves that they have already started to show potential and surely will have great societal impact, with growing interest on the part of industry to commercialize and market BCI systems for medical and non-medical applications both in the shortcoming and long term; the scientific challenge is clearly present.

References

1. Walter, W.G.: Contingent negative variation: an electrical sign of sensorimotor association and expectancy in the human brain. Nature 203, 380–384 (1964)
2. Sutton, S., et al.: Evoked potential correlates of stimulus uncertainty. Science 150, 1187–1188 (1965)
3. Naatanen, R.: Processing negativity: an evoked-potential reflection of selective attention. Psychological Bulletin 92, 605–640 (1982)
4. Picton, T.W.: The P300 Wave of the Human Event-Related Potential. Journal of Clinical Neurophysiology 9, 456–479 (1992)
5. Johnson, R.: On the neutral generators of the P300 component of the event related potential. Journal of Psychophysiology
6. Farwell, L.A., Donchin, E.: Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. Electroencephalography and clinical Neurophysiology 70, 510–523 (1988)
7. Middendorf, M., et al.: Brain-computer interfaces based on the steady-state visual-evoked response. IEEE Trans. Rehab. Eng., 8, 211–214
8. Wolpaw, J.R., et al.: Brain-computer interfaces for communication and control. Clin. Neurophysiol. 113, 767–791 (2002)
9. Berger, H.: Uber das Elektrenkephalogramm des Menschen. J. Psychol. Neurol. 40, 160–179 (1930)
10. Pfurtschelle, G., Lopes da Silva, F.H.: Event-related EEG/MEG synchronization and desynchronization: basic principles. Journal of Clinical Neurophysiology 110, 1842–1857 (1999)
11. Pfurtscheller, G., Aranibar, A.: Event-related cortical desynchronization detected by power measurements of scalp EEG. Electroenceph. clin. Neurophysiol. 42, 817–826 (1977)
12. Kalcher, J., et al.: Graz brain-computer interface II: towards communication between humans and computers based on online classification of three different EEG patterns. Med. & Biol. Eng. & Comput., 34, 382–388 (1996)
13. Pfurtscheller, G., Neuper, C.: Motor imagery and direct brain-computer communication. Proceedings of the IEEE 89, 1123–1134 (2001)
14. Hinterberger, T., et al.: Brain areas activated in fMRI during self-regulation of slow cortical potentials (SCPs). Experimental Brain Research 152, 113–122 (2003)
15. McFarland, D.J., et al.: Spatial filter selection for EEG-based communication. Electroencephalography and clinical Neurophysiology 103, 386–394 (1997)
16. Naeem, M., et al.: Seperability of four-class motor imagery data using independent component analysis. Journal of Neural Engineering 3, 208–216 (2006)
17. Bernat, E.M., Williams, W.J., Gehring, W.J.: Decomposing ERP time-frequency energy using PCA. Clinical Neurophysiology 116, 1314–1334 (2005)
18. Kachenoura, A., et al.: Ica: a potential tool for bci systems. IEEE Signal Processing Magazine 25, 57–68 (2008)
19. Hyvärinen, A., Karhunen, J., Oja, E.: Independent Component Analysis. Wiley, New York (2001)
20. Lee, T.W., Girolami, M., Sejnowski, T.J.: Independent component analysis using an extended infomax algorithm for mixed sub-gaussian and super-gaussian sources. Neural Computation 11, 417–441 (1999)

21. Jung, T.P., et al.: Removal of eye activity artifacts from visual event related potentials in normal and clinical subjects. Clinical Neurophysiology 111, 1745–1758 (2000)
22. Piccione, F., et al.: P300-based brain computer interface: Reliability and performance in healthy and paralysed participants. Clinical Neurophysiology (2006)
23. Xu, N., et al.: BCI competition 2003-data set iib: Enchancing p300 wave detection using ica-based subspace projections for BCI applications. IEEE Transactions on Biomedical Engineering 51, 1067–1072 (2004)
24. Wang, W., et al.: Lead selection for SSVEP-based brain-computer interface. In: EMBS 2004, 26th International Conference on Engineering in Medicine and Biology Society, September 2004, pp. 4507–4510 (2004)
25. Qin, L., Ding, L., He, B.: Motor imagery classification by means of source analysis for brain-computer interface applications. Neural Engineering 1, 135–141 (2004)
26. Koles, Z.J., Lazar, M.S., Zhou, S.Z.: Spatial patterns underlying population differences in the Backgorund EEG/. Brain Topography 2, 275–284 (1990)
27. Muller-Gerking, J., Pfurtscheller, G., Flyvbjerg, H.: Designing optimal spatial filters for single-trial EEG classification in a movement task. Clinical Neurophysiology, 110, 787–798 (1999)
28. Ramoser, H., Muller-Gerking, J., Pfurtscheller, G.: Optimal Spatial Filtering of Single Trial EEG During Imagined Hand Movement. IEEE Transactions on Rehabilitation Engineering 8, 441–446 (2000)
29. Townsend, G., Graimann, B., Pfurtscheller, G.: A comparison of common spatial patterns with complex band power features in a four-class BCI experiment. IEEE Transactions on Biomedical Engineering 53, 642–651 (2006)
30. Fischer, R.A.: The Use of Multiple Measurements in Taxonomic Problems. Annals of Eugenics 7, 179–188 (1936)
31. Muller-Putz, G.R., et al.: Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components. Journal of Neural Engineering 2, 123–130 (2005)
32. Geng, T., et al.: A Novel Design of 4-Class BCI Using Two Binary Classifiers and Parallel Mental Tasks. Computational Intelligence and Neuroscience (2008)
33. Burges, C.J.: A tutorial on support vector machines for pattern recongition. Data Mining Knowledge Discovery 2, 121–176 (1998)
34. Burmeister, J.J., Gerchardt, G.A.: Neurochemical arrays. In: Grimes, C., Dickey, E., Pishko, M.V. (eds.) Encyclopedia of sensors, vol. 6, p. 525. American Scientific Publishers, Stevenson Ranch (2006)
35. BrainGate System, `http://www.braingate2.org/` (Cited: September 7, 2009)
36. Coulson, S., et al.: Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. Electroencephalography and Clinical Neurophysiology 62, 203–208 (1985)
37. Wang, Y., et al.: Brain-Computer Interfaces Based on Visual Evoked Potentials. IEEE Engineering in Medicine and Biology Magazine 27, 64–71 (2008)
38. Silberstein, R.B., et al.: Steady state visually evoked potential (SSVEP) topography in a graded working memory task. International Journal of Psychophysiology 42, 219–232 (2001)
39. Donchin, E., Spencer, K.M., Wijesinghe, R.: The Mental Prosthesis: Assessing the Speed of a P300-Based Brain–Computer Interface. IEEE Trans. Rehab. Eng. 8 (June 2, 2000)
40. Birbaumer, N., et al.: Slow potentials of the cerebral cortex and behavior. Physiology Review 70, 1–40 (1990)

41. Kaper, M., et al.: BCI Competition 2003 - Data Set IIb: Support Vector Machines for the P300 Speller Paradigm. IEEE Transactions on Biomedical Engineering 51 (June 6, 2004)
42. Blankertz, B., et al.: The BCI Competition 2003: Progress and perspectives in detection and discrimination of EEG single trials. IEEE Transactions Biomedical Engineering 51, 1044–1051 (2004)
43. Schmidt, E.M.: Single neuron recording from motor cortex as a possible source of signals for control of external devices. Annals of Biomedical Engineering 8, 339–349
44. Principe, J.C., McFarland, D.J.: BMI/BCI modeling and signal processing. In: Berger, T.W., et al. (eds.) Brain-computer interfaces: an international assessment of reserarch and development trends, pp. 47–64. Springer Science + Business Media B.V., Berlin (2008)

Author Index