16TH ANNUAL EUROMEDIA CONFERENCE

2011

LONDON, UNITED KINGDOM

APRIL 18-20, 2011

Organized by

ETI

BITE
British Institute of Technology and Ecommerce

Sponsored by

EUROSIS

BITE

University of Skovde

Higher Technological Institute
Tenth of Ramadan City

Ghent University

TU Delft

HOSTED BY

BITE
The British Institute of Technology and Ecommerce

EUROMEDIA'2011

FEATURING

SIXTEENTH ANNUAL SCIENTIFIC CONFERENCE
ON WEB TECHNOLOGY, NEW MEDIA
COMMUNICATIONS AND TELEMATICS THEORY
METHODS, TOOLS AND APPLICATIONS

Hasan Al-Saedy

APRIL 18-20, 2011
LONDON, UNITED KINGDOM

A Publication of EUROSIS-ETI

Printed in Ghent, Belgium

EXECUTIVE EDITOR

PHILIPPE GERIL
(BELGIUM)

Editor

Professor Hasan Al-Saedy
The British Institute of Technology and Ecommerce, London, UK

Local Programme Committee

Professor Waheed Khan, BITE, London, UK
Professor Hasan Al-Saedy, BITE, London, UK
Dr. Peter Robinson, BITE, London, UK
Mr.Syed Majid, BITE, London, UK
Mr. Asheesh Rajouria, BITE, London, UK
Dr Joy Joseph, BITE, London, UK

International Programme Committee

WEBTEC Programme Committee

Ass. Prof. António Carlos Abelha, Universidade do Minho, Braga, Portugal
Boguslaw Butrylo, Bialystok Technical University, Bialystok, Poland
Khalil Drira, LAAS-CNRS, Université de Toulouse, Toulouse Cedex, France
Markus Koch, Orga Systems GmbH, Paderborn, Germany
Jens Lichtenberg, Ohio University, Athens, USA
Prof. Dr. Joern Loviscach, Fachhochshule Bielefeld, Bielefeld, Germany
Assoc. Prof. Wenji Mao, Chinese Academy of Sciences, Beijing, China P.R.
Lorenzo Motta, Ansaldo Segnalamento Ferroviaro s.p.a. Genova, Italy
Dr. H. Joachim Nern, Aspasia Knowledge Systems, Dusseldorf, Germany
Dr. Carlos E. Palau, Universidad Politecnica de Valencia, Valencia, Spain
Dr Stephan Reiff-Marganiec, University of Leicester, Leicester, UK
Prof. Paola Salomoni, Universita di Bologna, Bologna, Italy

MEDIATEC Programme Committee

Track Chair: Dr. Ignazio Infantino, Phd, ICAR-CNR, Palermo, Italy
IPC
Ass. Prof. Ali Arya, Carleton University, Ottawa, Ontario, Canada
Prof. Helena Barbas, CENTRIA-UNL, Lisbon, Portugal
Dr. Fernando Boronat Segui, Universidad Politecnica de Valencia, Gran de Gandia, Spain
Assoc. Prof. Vincent Charvillat, IRIT-ENSEEIHT, Toulouse cedex, France
Dr Jonathan Dukes, O'Reilly Institute, Trinity College, Dublin 2, Ireland
Jehan Francois Paris, University of Houston, Houston, USA
Florin Pop PhD, University POLITEHNICA of Bucharest, Bucharest, Romania
Prof. Marco Roccetti, Universita' di Bologna, Bologna, Italy
Dr. Leon Rothkrantz, Delft University of Technology, Delft, The Netherlands
Dr. Leonid Smalov, Coventry University, Coventry, United Kingdom

COMTEC Programme Committee

Prof. Dr. Marwan Al-Akaidi, De Montfort University, Leicester, United Kingdom
Christos Bouras, University of Patras, Patras, Greece
Boguslaw Butrylo, Bialystok Technical University, Bialystok, Poland
Sophie Chabridon, Institut Télécom, Télécom SudParis, Evry Cedex, France
Prof. Chris Guy, The University of Reading, Reading, United Kingdom
Ph. D. Oryal Tanir, Bell Canada, Montreal, Canada

INTERNATIONAL PROGRAMME COMMITTEE

APTEC Programme Committee

Dr. Juan Carlos Guerri Cebollada, Universidad Politecnica de Valencia, Valencia, Spain
Paulo Maciel, UFPE, Recife, Brazil
Assoc. Prof. Viorel Nicolau, "Dunarea de Jos" University of Galati, Romania
Dr.ir. Johan Opsommer, Belgacom - BUS, Brussels, Belgium
Prof. Jeanne Schreurs. Hasselt University, Diepenbeek, Belgium
Ass. Prof. Ramiro Velazquez, Universidad Panamericana, Aguascalientes, Mexico

E-TEC Programme Committee

Ciprian Dobre PhD, University Politehnica Bucharest, Bucharest, Romania
Jose Machado, University do Minho, Braga, Portugal
Paulo Novais, University do Minho, Braga, Portugal
Prof. Selwyn Piramuthu, University of Florida, Gainesville, FL, USA

Workshops

Cloud Computing

Prof. Dr. Jan Broeckhove, RUCA-UA, Antwerp, Belgium

Knowledge Management, E-Business management and E-Mobility

General Chair

Prof. Ricardo Chalmeta, Universidad Jaume I, Castellon, Spain
Jeanne Schreurs, Hasselt University, Diepenbeek, Belgium

In-Car Applications

Assoc Prof.Petr Hanacek, Brno University of Technology, Brno, Czech Republic

Virtual Reality Applications

Pieter Jorissen, Karel de Grote-Hogeschool, Hoboken, Belgium
Prof. Marcos A Rodrigues, CCRC Sheffield Hallam University, Sheffield, UK

INTERNATIONAL PROGRAMME COMMITTEE

Medical Imaging Systems

General Chair

Joao Manuel R. S. Tavares, FEUP, University of Porto, Porto, Portugal

General Co-Chair

Daniela Iacoviello, Department of Computer and System Sciences "A.Ruberti", Sapienza University of Rome, Italy

International Programme Committee

Adélia Sequeira, Instituto Superior Técnico, Portugal
Alberto De Santis, Università degli Studi di Roma "La Sapienza", Italy
Aledir Silveira Pereira São Paulo State University, Brazil
Ana Mafalda Reis, University of Porto, Portugal
Anton Vernet, University Rovira i Virgili, Spain
Arrate Muñoz Barrutia, University of Navarra, Spain
Begoña Calvo Calzada, University of Zaragoza, Spain
Bernard Gosselin, Faculte Polytechnique de Mons, Belgium
Christos E. Constantinou, Stanford University School of Medicine, USA
Dinggang Shen, UNC-CH School of Medicine, USA
Emmanuel A. Audenaert, Ghent University Hospital, Belgium
Enrique Alegre Gutiérrez, University of León, Spain
Fiorella Sgallari, University of Bologna, Italy
Jorge M. G. Barbosa, University of Porto, Portugal
Lyuba Alboul, Sheffield Hallam University, UK
Mahmoud El-Sakka, The University of Western Ontario London, Canada
Manuel González Hidalgo, Balearic Islands University, Spain
Maria Elizete Kunkel, Universität Ulm, Germany
Maria Petrou, Imperial College London, UK
Miguel Velhote Correia, University of Porto, Portugal
Paola Lecca, The Microsoft Research - University of Trento, Italy
Petia Radeva, Autonomous University of Barcelona, Spain
Renato Natal Jorge, University of Porto, Portugal
Sabina Tangaro, National Institute of Nuclear Physics, Italy
Teresa Mascarenhas, University of Porto, Portugal
Vassili Kovalev, University of Surrey, UK
Yongjie Zhang, Carnegie Mellon University, USA
Zeyun Yu, University of Wisconsin at Milwaukee, USA

Cybersecurity

Fernando de la Cuadra y de Colmenares, ONTINET.COM

Computer Graphics

Roberto de Beauclair Seixas, IMPA, Rio de Janeiro, Brazil
Philippe Geril, ETI, Ostend, Belgium

Student Track

Goreti Marreiros, Instituto Superior de Engenharia, Porto, Portugal

All articles published in these Proceedings have been peer reviewed

EUROSIS-ETI Publications are ISI-Thomson and INSPEC referenced
Legal Repository: Koninklijke Bibliotheek van België, Keizerslaan 4, 1000 Brussels, Belgium

For permission to publish a complete paper write EUROSIS, c/o Philippe Geril, ETI Executive Director, Greenbridge NV, Wetenschapspark 1, Plassendale 1, B-8400 Ostend Belgium

EUROSIS is a Division of ETI Bvba, The European Technology Institute, Torhoutsesteenweg 162, Box 4, B-8400 Ostend, Belgium

Printed in Belgium by Reproduct NV, Ghent, Belgium
Cover Design by Grafisch Bedrijf Lammaing, Ostend, Belgium

EUROSIS-ETI Publication

ISBN: 978-90-77381-62-5
EAN : 978-90-77381-62-5

Preface

Between the 18th and 20th of April 2011, the British Institute of Technology and E-Commerce (BITE) will host the following EUROSIS technical conferences EUROMEDIA'2011-ECEC'2011 and FUBUTEC'2011 in its Stratford campus in London. From the point of view of BITE this conference will be a landmark occasion in which more than fifty scientists from all over the world come together to discuss innovative scientific and engineering issues, ideas, concepts and patents.

The main streams of the research papers are on:

1- Complex System Simulation
2- Business Intelligence
3- Information Management and Security
4- Decision Support Tools
5- AI Applied to Concurrent Applications
6- Factory Planning and Control for Small and Medium Size Enterprises
7- Media Content Management
8- Data Transfer Architectures
9- Video on the Go
10- Media Interpretation System

In addition to the above mentioned streams, four significant key note speakers will participate in the conference to deliver lectures, discuss issues and chair sessions. The keynote speakers are:

1- Professor Fred Piper of Royal Holloway College, University of London who will discuss the current issues in security science and technologies. The title of his lecture is:
'Information Security: Finding the right balance'
2- Professor Des Mapps from University of Plymouth, speaking on current issues in data storage technologies. the title of his lecture is:
'Type and Limitations of Data Storage Hardware in Supporting Multimedia Platforms'
3- Professor Talib Alukaidey from Hertfordshire University will discuss advances in Networking Technologies. The title of his lecture is:
'The Socio-economic impact of ultra high bandwidth of internet network by 2020'
4- Professor Imed Kacem from Paul Verlaine University, Metz, France, will talk about advances in business simulation and the title of his lecture is:
'Simulation approaches for optimizing in business and services systems'

Preface

In addition to the above mentioned delegates a large number of BITE teaching staff and postgraduate students will participate in the conference sessions. The postgraduate students expected to participate in conference are from the following courses:

1- Innovative Technology
2- Security Technology
3- Network Technology
4- Animation Technology

Professor Hasan Al-Saedy
Professor of Information Security
British Institute of Technology and E-Commerce

CONTENTS

CONTENTS

MEDIA INTERPRETATION SYSTEMS

SCIENTIFIC PROGRAMME

MEDIA CONTENT MANAGEMENT

NARRATIVE BIFURCATION, CLOUD COMPUTING INTERFACES AND HITCHCOCK

Helena Barbas
IEMo and Department of Portuguese Studies
Faculdade de Ciências Sociais e Humanas – Universidade Nova de Lisboa
1069-061 Lisbon,
Portugal
E-mail: hebarbas@fcsh.unl.pt

KEYWORDS
Storytelling; narratology; attention span; hyperfiction; serious games; digital humanities; information retrieval;

ABSTRACT

From a narratological perspective, this paper aims to address the theoretical issues concerning the functioning of the so called «narrative bifurcation» in data presentation and information retrieval. Its use in cyberspace calls for a reassessment as a storytelling device. Films have shown its fundamental role for the creation of suspense. Interactive fiction and games have unveiled the possibility of plots with multiple choices, giving continuity to cinema split-screen experiences. Using practical examples, this paper will show how this storytelling tool returns to its primitive form and ends up by conditioning cloud computing interface design.

INTRODUCTION

Narrative theories have allowed the interpretation of traditional narrative fiction to be based on some quite sound structural foundations. However, new discoveries in other sciences and the practices of hypermedia are providing grounds for updates and reformulations.

Most commonly applied to the mathematical study of dynamic systems, bifurcation occurs when a change made to a structure causes a sudden qualitative or topological alteration in its behaviour. In discourse, bifurcation is present in the choice of words, conditioned by the vocal tract (Benveniste 1974). The idea of bifurcation is reinforced by a binary logic mode of thinking. It should be reassessed in the light of the newer formal logics, namely the resultant from computer programming (three value or fuzzy logics). This could bridge one of the gaps between arts and sciences, since the latter, that relied on strict logical systems where things were black or white, can get closer to the former, hovering in a grey, fuzzy, intermediate area.

NARRATIVE BIFURCATION

Bifurcation is an outcome of the causal and chronological dimensions of all kinds of discourse. It is essential to the functioning of any kind of narrative. Here it will be adopted the definition of narrative already offered (Barbas 2010; Ryan 2001), and reinforced the idea that all discourse is a narrative. An account corresponds to the representation of at least one event; each event is given under the form of two propositions (sentences). The story results from the relationship established between those two sentences (by contiguity and consecution). In its core, this definition of narrative can be extrapolated to the act of offering any kind of information – being it scientific or other. It is this feature that will condition any kind of information retrieval, by any user, from any interface.

Bifurcation and simultaneity

In written or oral narrations simultaneous events are rendered sequentially. As we have seen (Barbas 2010; Monfort 2007) the order of the presentation can be crucial to the meaning of the story. In traditional narrative, bifurcation implies that the author will choose one of the events discarding the other, the outcome of a conflict; the user/reader will have to comply and accept it.

At plot level, this procedure can be used as a stylistic tool. The author can give fractions/sequences of each action one after the other, alternatively. This is done in written fiction, but the strategy is more evident in cinema or hyperfiction, and is the basic tool for the creation of suspense. In the wake of Hitchcock, Brewer and Lichtenstein (1982) have claimed that it is possible to instigate affective states in the reader by arranging the temporal order of the events underlying a story world. Their theory explains that suspense could be evoked by presenting the events of a story chronologically to the reader. Yet, against Hitchcock – to whom the reader/spectator must be God, and know more than the characters – they suggest that surprise and curiosity could be caused by hiding a critical fact or event early in the story world and disclosing it later in the text.

BIFURCATION AND SUSPENSE

For Hitchcock: «On the screen, terror is induced by surprise; suspense by forewarning» (Gottlieb 1995). Suspense has been defined as a feeling (uncertainty) or an emotion (anxiety) concerning the outcome of certain actions – dramatic or not. The main problem is that the effect (resulting sensation/secondary) is being defined by the cause (play with time/primary expectations).

Suspense results from the manipulation of order and time of narrative sequences. Says Hitchcock: «A light hearted comedy played slowly may produce the sense of impending doom, just as a too brightly acted drama might never give an atmosphere of tragedy». Suspense is achieved by the order chosen to present the actions (montage), or by the delay in the resolution of the bifurcation (barthesian catalysis).

The play with order has already been addressed. The problem of duration/time has to include a factor external to

the narrative. To be effective, suspense has to consider the dynamics of the user attention span – says Hitchcock: «How many people realize, I wonder, that we do aim at moods in our films? We call it 'Tempo' and by paying careful attention to the speed with which we act our little plays we do attempt to guide the observing minds into the right moods.» (Gotlieb 1965:167).

'Tempo' and the attention span

The Hitchcokian 'Tempo' could be a synonym to the attention span: the amount of time that anyone can concentrate on a task. Studies on attention span (Middendorf and Kalish, 1996, i.e.) are usually centered on student performances. But there is a general agreement that human beings have, at least, two types of attention, short and long term. The first, is a focused, brief response to a stimulus, that may be as fast as 8 seconds. The latter, is a sustained effort that allows the production of consistent results on a task over time. The range – for an healthy adult – reaches the maximum of 20 minutes at a time. This span builds up in a curb, that can be renewed, and that allows the individual to concentrate in longer tasks – being it study, watching a movie or browsing the web. Hitchcock builds his film sequences playing with this two span factors: the greatest amount of information in the shortest time (4/12 seconds) up to the climax in circa 12 minutes.

Like a goldfish

The controversies surrounding these attention spans have been transferred to the new media. As it happened with books, authors believe that the human attention scope is decreasing as modern technology, especially television, increases.

Internet browsing is charged with a similar effect (Carr 2010) because it enables users to easily move from one page to the other: «Navigating linked documents, it turned out, entails a lot of mental calisthenics – evaluating hyperlinks, deciding whether to click, adjusting to different formats — that are extraneous to the process of reading.» A 2002 article from BBC News, still quoted (Ebert 2010), forebodes that we will all end up with the attention span of a goldfish.

User studies – including tracked eye movement ones – show that people start to read faster and less thoroughly as soon as they go online (Carr 2010). The argument is that «each glance breaks our concentration and burdens our working memory, [and] the cognitive penalty can be severe». The penalty is amplified by what brain scientists call switching costs. Every time we shift our attention, the brain has to reorient itself, further taxing our mental resources. Yet, it seems that every medium develops some cognitive skills at the expense of others (Greenfield 2009).

These studies do not specify which kind of attention was being measured – shot or long term. And the issue was to determine what or how can make the users «stick» to a website. It seems that web surfers show a «maddening unwillingness to stay put». Ever on the lookout for engaging content, most online viewers spend less than 60 seconds at an average site.

Yet, internet should not affect the attention span anymore than a book with footnotes, endnotes, accreditations, bibliography and an index. Humans use their concentration in accordance with the tasks they are executing. Browsing should naturally require a short term attention span – the same as the necessary to peruse a telephone directory – and the switch to long term attention will logically occur when the effort to obtain material/information becomes relevant.

Programming Suspense

Besides the above differences in human attention span, user studies should also consider as a useful mesure the 'Tempo' /speed of the information provided.

The technique of postponing story resolution/delaying bifurcation is easy to computerize. The MINSTREL system, i.e., includes additional events that detail the protagonist's struggles in between the story's climax and its resolution (Yacine 2011). However, experiences with algorithms to program 'suspense' have not succeed: «While the results of this study do show that Suspender was effective in generating suspenseful stories, the design of the experiment does not allow us to point conclusively at a single reason for this effectiveness» (Cheong et al 2008).

BIFURCATION AND MONTAGE

Editing is the signature of an author, his style. It plays the main role in Hitchcock's creation of suspense. Each juxtaposition of shots is carefully prepared to achieve its utmost impact. The famous shower scene in *Psycho* (1960) uses montage to hide the violence. The knife stabbing the body is never shown; the illusion of violence is created through the editing speed. In *The Man who knew Too Much* (1956) climax is achieved in 12 minutes. There is a plot to assassinate a statesman at a concert, and the audience is immediately informed of the exact point in time when the crime will occur. «The very first thing that Hitchcock does is to establish the situation and indicate all who will be involved. This is done for story purposes, as well as to keep with his principle of constantly informing the audience, in an efficient 38-second sequence.» (Sly 2008). Following this cue to analyze *The Rope* (1948), Brandon and Philip strangle David Kentley in the first 3 minutes; the hiding of the corpse and the set up for dinner are ready at minute 11:36 – when a close-up shows the audience the poorly hidden rope.

FROM THIRD OPTION TO MULTILIEAR PLOT

In storytelling, the existence of a choice at the level of the middle function creates a binary opposition up to now considered impossible to surmount. This matter has been addressed by hyper fiction practitioners and theorists.

Chris Crawford (Crawford 2004) considers that more than two alternatives are irrelevant: «If the user has reached the climax of the story and must choose between leaving his girlfriend for the war and shirking his duty, having only two choices doesn't detract from the power of the interaction; it's difficult to imagine any other reasonable possibilities». Further on, affirms that the computer cannot offer the third option, because «the correct one would be obvious to the reader». He is right, but for the wrong reasons – the problem is that the reader can only follow one option at a time.

Emily Short contradicts him with her practice – since her first work, *Galatea* (2000), with multiple plots and endings: «Galatea has what I call a multilinear plot: unlike traditional IF, it has no single path to victory. Instead there are a large

number of endings, some more satisfactory than others, of which many could be considered "win" states» (Short 2008). In her collective work *Alabaster* (2008), a «fractured fairy tale» inspired by *Snow White*: «The Queen is certainly a witch – but her stepdaughter may be something even more horrible… There are some eighteen possible endings to this fairy tale. Some of them are even almost happy».

The multiplicity of plots – of «third options» – may allow the user to experience large segments of the game in different ways, on different plays. These possibilities have already been explored in games that may contain entire alternate scenes and segments (*Losing your grip*, Stephen Granade, 1998; *Scavenger*, Quintin Stone, 2003; *Heavy Rain*, Sony, 2010) or multiple endings (*Slouching Towards Bedlam*, Daniel Ravipinto and Star Foster, 2003). But in the end neither evades bifurcation.

LITERARY AND CINEMATIC EXPERIENCES

This polemics can have a retroaction effect on narratology. The two options alternative has been proved correct with the literary analyses of traditional linear narrative because it was the norm. Yet, there have been literary experiences of more than one option, and multiple endings hypothesis.

Theoretically, the problem is addressed in Umberto Eco's *Opera Aberta* (1979) – where the ending, corresponding to a total absence of viable alternatives, or alternatives with equal weight, are (supposedly) left to the reader. In practice, there have been experiments to surmount bifurcation and present the reader with multiple choices. This consists in offering, at plot structural level, three possible and equivalent ways of solving a conflict. As to the theatre, there is a very dramatic "third option" in the tragedy *Oedipus Rex* (III vv. 950-970) by Sophocles (c. 497/6-406/5 BC). Curiously, as will be done by Hitchcock, the suspense/solution to the dilemma, results from the fact that the audience has all the premises (three) and the tragic characters only two.

Playing with bifurcation is the leitmotiv of the first of the novels, *Life and Opinions of Tristram Shandy* (1759-1789), by Laurence Sterne (1713-1768). Amongst the most recent works of fiction (prior to the pc's boom) it must be referred *The French Lieutenant Woman* (1969) by John Fowles (1926-2005) presenting three alternative endings (not so well worked-out in the 1981 film directed by Karel Reisz with Harold Pinter's screenplay).

In cinema there is a recent inconspicuous film whose intrigue is completely built upon the hypothesis of a "third" hidden option, always prepared by the plot, but becoming visible only when the actual scene is enacted – *Dirty Rotten Scoundrels* (1988), directed by Franz Oz, with Michael Caine, Steve Martin, and Gleanne Headly, with script from Dale Launer, Stanley Shapiro Paul Henning.

All these examples prove, however, that, independently of the quantity of plots or endings available, the reader will only, and ever, be able to follow one of them at each time.

Simultaneity and split screen

So, simultaneity is the problem of dealing with two events occurring at the same moment inside the narrative. In cybertext, tv serials and cinema it can be enacted through the technique called «split-screen». This is the visible division of the screen in half, or presenting several simultaneous images. It intends to destroy the illusion that the screen's frame is a seamless view of reality, similar to that of the human eye.

This intercut technique of dividing the screen has its first example in *Are You There*? (1901) a 35mm, black and white silent movie directed by James Williamson.

Figure 1: James Williamson - *Are You There ?* (1901)

In *Suspense* (1913), a thriller directed by Phillips Smalley and Lois Weber, this technique is called the triptych – to show several actions occurring at the same time.

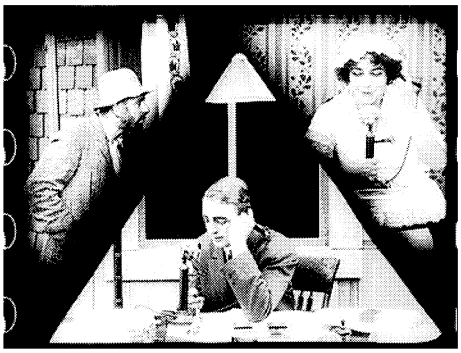


Figure 2: Phillips Smalley and Lois Weber - *Suspense* (1913)

The telephone as an pretext is also the theme of more recent and famous split-screen movies – from *Pillow Talk* (1959) directed by Michael Gordon, with Rock Hudson and Doris Day (Academy Award for Writing Original Screenplay) to *Eat, Pray, Love* (2009) directed by Ryan Murphy, with Julie Andrews.

From split-screen to mosaic

The intercut technique dividing the screen in two halves, or the montage technique allowing it to incorporate several shots of different events in one scene, have exploded with the arrival of digital video technology. The division is much easier to accomplish, and recent digital films and music videos have explored this possibility widely and in depth. It has almost become the norm: in tv serials, since *24* (2001), all have several scenes using this technique.

The split screen sequences – even when the screen is filled-up with images – eventually are reduced to only one picture, conventionally occupying solo the screen. The same happens with the fragmentary «mosaic-screen» (Branco 2008), producing similar effects. Both options fit into the use of multi-frame imagery; and the problem with both is related with the way humans look at an image: if the screen is

partitioned in more than four sections, the message becomes unreadable.

Also, the presence of multiple frames needs an attention focusing point. This can be achieved by making one frame a little larger than the others; by giving the dialogue/sound of only one of the scenes; or by playing a unique soundtrack. Whatever the choice, narrative unity has to be secured.

CLOUD COMPUTING AND INTERFACES

Cloud computing is the movement towards a network based environment, implying novel conceptual models for storing and sharing digital information (Terrenghi 2009; Pan 2011). The server is replacing the desktop. And the exhibition and visualization of resources are exploring the imagetic possibilities offered by split- and mosaic-screen, following the latest trends.

Cool iris - Cooliris (http://www.colliris.com) is a free downloadable plug-in that turns any browser into a super media browser at the click of a mouse. Images are presented as a waving gallery, and the user can flip through them as in a catalogue.

Youtube Doubler and **You3b** – provide the possibility of playing two or three videos at the same time. **Youcube** (http://youcub.es/) by Aaron Meyers (2009) is an open source software that allows visitors to map six YouTube videos onto a cube that can be manipulated in 3D space. As the cube rotates, the corresponding video sounds fade in and out of audible range.

Microsoft has upgraded to **SilverLight 5**, promising to change the face of UI on the web. PivotViewer (http://www.microsoft.com/silverlight/pivotviewer/) is said to make it easier «to interact with **massive amounts** of data on the web in ways that are powerful, informative, and fun. By visualizing thousands of related items at once, users can see trends and patterns that would be hidden when looking at **one item at a time**», and: «the animations and natural transitions provide context and **prevent users from feeling overwhelmed** by **large quantities** of information» [bold is ours]. The propaganda hides a «denegation», users are overwhelmed by massive amounts of information, and can oly look at one item at a time. The presentation of information in multifaceted windows/screens is invading the web and will certainly change the user experience. However, the user will continue to act at her own pace.

CONCLUSION: BIFURCATION RETURNS

The web and computers are being invaded by new forms of presenting the information. However, quantity is not quality even when aesthetically presented. It is humanly impossible to look at more than one item at a time. So, from the screen side, narrative bifurcation issues are still at work, in presence of massive amounts and mosaics of information. From the user side, behaviour is conditioned by physique, actions and reactions, brain functioning and human attention scope. Even in the case of long term, the span will be focused only on one task at a time, and not for longer than 20 minutes, with a peak of attention around 13 minutes – just like in classes, in lectures or Hitchcock's movies.

Considering that the human user's brain will keep working at its own pace, cloud computing interfaces will have to comply with Hitchcock's 'Tempo'.

REFERENCES

Barbas, H. 2010. "Narrative Memory in Hyperfiction and Games". *GameOn'2010 - 11th International Conference on Intelligent Games and Simulation,* ed. Aladdyn Ayesh, Eti/Eurosis, De Monfort University, Leicester, U.K., 17-19.Nov., pp. 85- 91.

Barthes, R. 1966. "Une Introduction à L'Analyse Structurelle du Récit". *Communications 8*. Seuil, Paris, France.

Benveniste, É. 1974. "L'Homme dans la Langue". *Problèmes de Linguistique Générale, 2*. Gall., Paris, France. 195-238.

Brewer, W.F., Lichtenstein, E.H. 1982. "Stories Are to Entertain: A Structural-Affect Theory of Stories". *Journal of Pragmatics* 6(5-6), 437–486;

Carr, N. 2010. *The Shallows: What the Internet Is Doing to Our Brains.* W.W. Norton and Company, London.

Cheong, Y-G., Young, M., et al. 2008."Narrative Generation for Suspense: Modeling and - Evaluation. U. Spierling and N. Szilas (Eds.), Springer-Verlag, Berlin Heidelberg, pp. 144–155.

Crawford, C. 2004. *On Interactive Storytelling*. New Riders, Indianapolis. pp. 41

Ebert, R. 2010. "The Quest for *frisson*". *Chicago Sun Times* blog, May 29, 2010.

Eco, U. 1979. *L'Œuvre Ouverte*. Seuil, Paris.

Genette, G. 1966. "Fontières du Récit". *Communications 8*. Seuil. Paris, France.152-63.

Greimas, A. J. 1966. *Sémantique Structurale. Recherche de Méthode*. Larousse, Paris, France.

Gottlieb, S. (ed).1995. *Hitchcock on Hitchcock*. University of California Press. Berkely, Los Angeles, London.

Greenfield, P.M. 2009. "Social change and human development: A theory for the data". *ISSBD Bulletin,* No. 1, Serial No.55, 20-23.

Holtzman, S. 1997. *Digital Mosaics: The Aesthetics of Cyberspace*, Simon & Schuster, N. York.

Pan, Y., Blevis, E., 2011. "The Cloud". *Interactions*. Jan-Feb 2011, Vol.XVIII, nº.1, A.C.M. p.13-15;

Terrenghi, L., et al. 2009. "Elastic mobility: stretching interaction". *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '09, art.46 , 4 pp. ACM. New York.

WEB REFERENCES

Branco, S.D. 2008. "The Mosaic-Screen: Exploration and Definition". *Refractory*, 27.Dec.2008; http://blogs.arts.unimelb.edu.au/refractory/2008/12/27/the-mosaic-screen-exploration-and-definition-%E2%80%93-sergio-dias-branco/ [Jan.2011]

Middendorf, J. and Kalish, A. 1996. "The 'Change Up' in Lectures". *The National Teaching & Learning Forum*. Teaching Resources Center. Indiana University. Jan. Vol.5.Nº.2. http://www.ntlf.com/html/pi/9601/article1.htm [Jan.2011]

Yacine, 2011. *Minstrel Stories* - http://www.lvl31.com/minstrel-stories-volume-1/ [Jan.2011]

BIOGRAPHY

HELENA BARBAS (1951) is Professor-Lecturer of the Departament of Portuguese Studies – Faculdade de Ciências Sociais e Humanas – Universidade Nova de Lisboa; she is a researcher of IEMo and collaborator of CENTRIA. She holds a MA (1990) and a PhD (1998) in Comparative Literature – Literature and the Arts, and gained her "Habilitation" (2008) in Literature and Cyberarts. Research interests: Digital humanities, storytelling, hyperficton, avatars, serious games. URL: http://www.helenabarbas.net.

AN APPROACH TO LEARNING-BY DOING THROUGH USER CREATION OF LEARNING CONTENT

Danail Dochev, Gennady Agre
Institute of Information and Communication Technologies – Bulgarian Academy of Sciences /BAS/
Sofia, Bulgaria
E-mail: dochev@iinf.bas.bg, agre@iinf.bas.bg

Radoslav Pavlov
Institute of Mathematics and Informatics – BAS
Sofia, Bulgaria
E-mail: radko@cc.bas.bg

KEYWORDS
CAI, learning-by-doing, semantic technologies

ABSTRACT
The paper investigates the organisation of learning-by-doing activities through learner's authoring of analytical materials, considering some specifics of education in humanities. Semantic technologies are applied to guide and help the learner's actions in developing limited-sized dedicated collections of multimedia objects, adequate to pre-assigned learning tasks and then in comparing the selected objects for the needs of performed analysis. The discussed framework for a TEL environment targets concrete humanitarian domain - Bulgarian Iconography with educational uses in a set of disciplines like iconography, arts, history, culture studies, theology, etc. The paper contains an example with structured formulation of a given learning task, and its formalisation in form of queries for the environment to help in the collection development steps and in the evaluation of the adequacy of the selected by the learner representative subset of objects.

INTRODUCTION

In the educational theory of constructivism (Driscoll, 2000) whose academic significance and practical value are generally acknowledged, it is universally recognized that the learner needs an active learning style, different from passive reading and memorizing. The learners have to participate actively in learning situations in order to "construct" their own understanding of the subject. In case of education in humanities a set of specific features have to be considered: a/ Big volumes of not explicitly structured knowledge; b/ Domain significant, yet differing theoretical models, which can not be generalized in a common framework; c/ Different, even contradictory interpretations of phenomena, which can not and should not be neglected in the learning process; d/ More significant impact of linguistic, cultural and subjective factors on the understanding and explanation of phenomena; e/ Conceptual theoretic frameworks, built on not fully defined concepts and notions, hampering the computer modelling.

The present paper investigates learning-by-doing through learner's authoring of analytical materials, facilitated by applying semantic technologies to support the learners in the access and filtration of necessary multimedia objects to be analysed during the authoring process, as well as in the materials evaluation. The presented work is organised under current national research project SINUS "Semantic Technologies for Web Services and Technology Enhanced Learning" (Dochev and Agre 2009).

LEARNING-BY-DOING THROUGH AUTHORING

Learning setting and objectives

The investigations in the project SINUS aim to support the analytical and to a certain extent the interpretative skills of the learners in a given humanitarian field by developing an environment, permitting to build-up appropriate task-focused presentation of annotated digital resources to the users. These resources are intended for use in the authoring of analytical materials in defined learning situations. The participation of the learners in:

- semi-structured navigation on appropriate information objects (the structure is offered by domain ontology);
- writing of analytical materials on selected objects (with offered guidance by the system)

will guarantee their active involvement in the learning process in constructivist sense.

The project SINUS aims to a specialised learning environment, supporting the learners to create specific learning materials by intensive use of multimedia digital libraries /DL/. With these 'learning-by-authoring' activities the following learning goals are pursued: a/ Improving, making more precise, consolidating and extending learner's specific domain knowledge; b/ Developing the analytical skills of the learners and facilitating their application; c/ Mastering professional DL usage by the learners.

The SINUS environment supports specific learning tasks - development of educational scholarly essays/course theses/ projects for pre-assigned by the teacher analyses of the objects under study. The result is a multimedia document combining a specific collection of DL multimedia objects, and textual analytical essay (e.g. analysis of certain object characteristics). The developed collection serves as a base for performing the necessary analysis and as an illustration of theses in the analytical essay. A project development consists of three steps: 1/ Constructing limited-sized dedicated multimedia collection by selecting appropriate material from DL with semantically annotated resources; 2/ Analysis of the selected collection by comparison and debate of certain objects characteristics. The analysis may require modification/enrichment of the developed collection. 3/ Development of the analytical essay as a multimedia document.

The environment guides and consults the learners on the base of its built-in knowledge (implemented by use of semantic technologies): A/ Domain knowledge about the subject under study, presenting its concepts and relations, significant from the viewpoint of the learning process. It is organized in a set of domain ontologies. The information and learning materials in the repositories are semantically annotated in the terms of these ontologies. B/ Pedagogical knowledge, reflecting the teacher/expert mental picture about the content, structure and steps to create good analytical essays, based on sufficiently rich and various illustrative materials.

Both types of knowledge have to consider the normal shortage, inaccuracy and even incorrectness of the initial learners' knowledge for the domain and also for the accessible materials and available information support.

Development of a collection

From the learners' viewpoint the development of a multimedia collection is a cyclic process of two steps: 1/ Formulation of a query, describing certain desired characteristics of the collection; 2/ Selection of representative objects from the set satisfying the query. From the environment side this is a cyclic process of performing the following functions: a/ Finding a set of objects, satisfying the query; b/ Sorting the result according implicit or explicit criteria; c/ Visualisation of sorted objects with different descriptive details according implicit or explicit criteria; d/ Storing the user choice (subset of the search results) in appropriate form for further processing.

The environment aims to facilitate the learner's actions by intelligent (i.e. guided by additional built-in knowledge) execution of system functions, listed above. The additional knowledge is in a sense pedagogic knowledge and determines concrete criteria (parameters) for the system functions. It permits to control the correctness of the learners' actions when performing the sequence of operations to solve the task. This type of knowledge may be used to:

- Help the student to formulate concrete query, by using friendly interface with appropriate ontological terms;
- Explain some unknown/unclear terms and their relations with familiar concepts;
- Check the query result correctness (e.g. by comparison against an exemplary query for given task);
- Facilitate the student's selection of representative objects (by different modes of visualization);
- Check the adequacy of the selected objects against teacher's criteria– number, variety, area coverage etc.

In reality the relationship between learning tasks and possible paths through the space of learning resources is 'many to many': one task can be achieved by following several paths and vice versa. SINUS environment considers this by allowing the student to perform the assigned learning task with the system help in 3 modes: 1/ Independently from the offered plan by the teacher, using only the semantic search and explanation facilities of the environment. 2/ Using the recommendations in the teacher's plan and the system help. The environment does not monitor the execution steps, but only evaluates the adequacy of the developed collection against the learning task. 3/ Using step-by-step recommendations of the plan and full system help, including monitoring and evaluation of queries with feedback to the learner at each step.

Collection analysis

The collection analysis begins with comparison of given collection objects in order to determine their common or differing characteristics. The following operations will contribute to this comparison:

- Grouping of the collection elements according to given value/s of their characteristic/s;
- Partitioning on subgroups according to given characteristic or set of characteristics;
- Finding partitioning according to common value/s of characteristic/s (the values are not pre-defined);
- Ordering objects according a characteristic value (e.g. chronologically);
- Registering similarities and differences of the selected elements according pre-defined characteristics.

To obtain information support during the initial steps of collection analysis the learner should formulate correct queries to the environment and register the observations on the results. To ensure this support in addition to sorting and visualisation functions the learning environment has to execute commands like Group, Order, Compare, Register. The specific pedagogic-type knowledge should determine parameters for the internal functions, realising the commands.

In SINUS environment the evaluation of intermediate and final results is focused not on assessment phases, but on monitoring the process and helping the learners to create projects according to teachers' mental picture what is a good analytical essay (reflected in the built-in knowledge). The environment may evaluate if the developed task-focused collection of multimedia objects contains sufficiently rich illustrative material to back-up the analyses (e.g. checking minimal number of objects, sufficient coverage of: iconographic schools, time periods, diversity of analysed characteristics etc.).

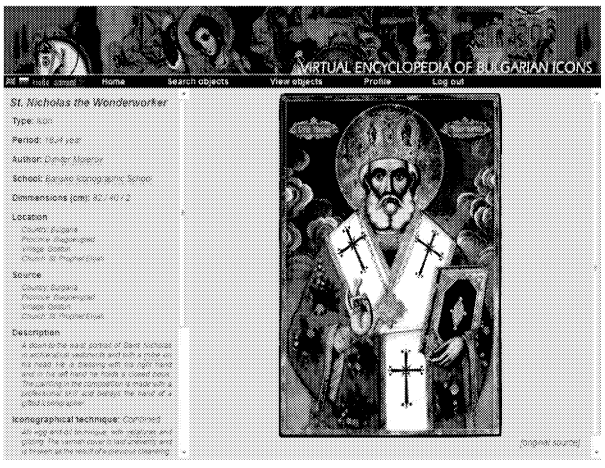


Figure 1: An object from the "Virtual Encyclopaedia of the Bulgarian Iconography"

AN APPLICATION EXAMPLE

The discussed framework for a TEL environment, facilitating learning-by-authoring in specific learning settings, is under experimental realisation for a concrete humanitarian domain - Bulgarian Iconography, which has educational uses in a set of disciplines like iconography, arts, history, culture studies, theology, etc. It is experimented on source materials from the multimedia digital library "Virtual Encyclopaedia of the Bulgarian Iconography" (Figure 1) (Pavlova-Draganova et al. 2007). The semantic annotation and search of digital objects is based on the domain Ontology for Bulgarian Iconographic Objects (Staykova and Dochev 2009).

The example below presents structured formulation of a concrete learning task, extracted from the SINUS project use-case scenarios (Table 1).

Table 1: Text representation of a learning task and recommendations

Task	Make critical art analysis of the chronological development of the iconographic image of Jesus Christ in the Bulgarian iconographic schools.
Step	***1. Select collection of objects for the analysis.***
Recommen-dations	The selected objects have to satisfy the following criteria:
Basic:	Select at least 6 iconographical objects with the person of Jesus Christ in compositions with one main figure. All iconographical objects have to be in good current condition.
Optional:	At least one object from eminent author or founder of iconographic school. At least one primitive iconographical object and at least one iconographical object from the period of Bulgarian renaissance.
Step	***2. Make analysis of the collection.***
Recommen-dations	Examine the selected objects, comparing: • the cloths, gestures, proportions of the person of Jesus Christ; • objects, other persons, Christian symbols; • background, other elements around the image of Jesus Christ. Search for changes – appearance or lack of elements (objects, symbols, persons), changes in background, clothes etc.
Step	***3. Register the results of the critical art analysis as a project.***
Recommen-dations	The project to be formed as multimedia document containing the selected iconographic images together with explanatory text before/after each image.

The formulation is further formalised in form of internal queries (presented in pseudo-language) for the environment to help in the collection development and evaluation of the adequacy of the selected by the learner representative subset of objects for comparison of their significant characteristics.

1. Query for search in the digital library according the recommendations (main criterion):

```
FIND  all ?Objects On_criterion:
      class_of(?Objects)= iconographic_object
AND contains_person(?Objects)=
     Jesus_Christ
AND participates(Jesus_Christ) =
     ?Composition
AND class_of(?Composition) =
     one_figure_composition
AND current_status(?Objects) = good
```

2. Query for sorting the selected objects:

```
ORDER  ?Objects On_criterion:
     iconographic_school (?Objects) =
            ?Iconographic_school
     AND author(?Objects) = ?Author
     AND (school_founder(?Author) = yes
       OR eminent_author(?Author) = yes)
```

3. Query for visualisation:

```
PRESENT   ?Objects On_criterion:
     composition(?Objects) =
            ?Composition
     AND class_of(?Composition) =
     one_figure_composition
     AND iconographic_school((?Objects) =
            ?Iconographic_school
     AND school_founder(?Author) = yes
     AND eminent_author(?Author) = yes
     AND primitive(?Object) = yes
     AND renaissance_object(?Object) = yes
```

4. Query for evaluation the adequacy of the selected by the learner representative subset of objects:

```
EVALUATE   ?Selected_objects =
     part_of(?Objects) On_criterion:
/* criterion for search */
class_of(?Selected_objects) =
            iconographic_object
AND contains_person(?Selected_objects)
            = Jesus_Christ
AND participates(Jesus_Christ) =
     ?Composition
AND class_of(?Composition) =
            one_figure_composition
AND current_status(?Selected_objects)
     = good
/* criterion for selection */
 AND ?Quantity(?Selected_objects) ≥ 6
 AND exists(?Object1) =
            part_of(?Selected_objects)
 AND author(?Object1) = ?Author
 AND (school_founder(?Author) = yes
     OR eminent_author(?Author) = yes)
 AND ?Quantity(?Object1) ≥ 1
 AND exists(?Object2) =
            part_of(?Selected_objects)
 AND primitive(?Object2) = yes
 AND renaissance_object(?Object2) = yes
 AND ?Quantity(?Object2) ≥ 1
 AND exists(?Object3) =
            part_of(?Selected_objects)
 AND renaissance_object(?Object3) = yes
 AND ?Quantity(?Object3) ≥ 1
```

CONCLUSIONS

An experimental learning environment for learning-by-doing activities through learners' authoring of analytical materials in specific learning settings is under development. It concerns the area of Bulgarian iconography, studied in a set of humanitarian disciplines. The current investigations and experiments led to the following possible future work directions: facilitating the semantic search queries by use of multilingual ontology-backed terminological lexicons; use of text processing and data mining techniques in formalisation of learning task descriptions to monitor and support collection development with rich enough illustrative material; use of similar techniques to monitor and support the preparation of analytical essays with appropriate argumentation. (e.g. checking the availability of necessary basic concepts in the text and the possibilities that they are semantically related).

ACKNOWLEDGEMENTS

The work on this paper was funded partially by the Bulgarian NSF project D-002-189 SINUS "Semantic Technologies for Web Services and Technology Enhanced Learning".

REFERENCES

Dochev D. and Agre G., 2009. Towards Semantic Web Enhanced Learning. In *Proc. of the Int. Conf. on Knowledge Management and Information Sharing*, Funchal, Madeira, pp. 212-217.

Driscoll, M., 2000. *Psychology of Learning for Instruction*. Needham Heights, MA, Allyn & Bacon.

Pavlova-Draganova, L., Georgiev, V., Draganov, L., 2007. Virtual Encyclopaedia of Bulgarian Iconography, *"Information Technologies and Knowledge",* vol.1, №3, pp. 267-271.

Staykova K., Dochev D., 2009. Ontology "Bulgarian Iconographical Objects"- Creation and Experimental Use, *"Cybernetics and Information Technologies",* Vol. 9 (2009), №1, pp. 25-37.

GAP4APE: a client-side system to enhance accessibility in the Web 2.0 era

Silvia Mirri
Department of Computer Science
University of Bologna
Via Mura Anteo Zamboni 7
40127 Bologna (BO), Italy
silvia.mirri@unibo.it

Paola Salomoni
Department of Computer Science
University of Bologna
Via Mura Anteo Zamboni 7
40127 Bologna (BO), Italy
paola.salomoni@unibo.it

Giorgia Cucchiarini
Corso di Laurea in Scienze dell'Informazione
University of Bologna
Via Sacchi, 3
47521 Cesena (FC), Italy
giorgia.cucchiarini@studio.unibo.it

ABSTRACT

Social networking systems have radically changed personal communication by providing new and interesting opportunities both in leisure and in business. Despite their apparent universal appeal, these systems are effectively leaving out a part of users with disabilities, who have difficulties in register, join and participate in the main on-line communities. Facebook represents a critical case: beside some accessibility improvement provided by this social network service, its accessibility represents an issue yet. This paper presents an augmented browsing system, which allows users with disabilities to specify their needs and preferences about Web pages presentation, directly by means of the browser interface. On the basis of the declared settings, the system automatically transcodes the page content (both static and dynamic), producing new pages adapted to the user's needs. The system, based on the widespread Web browser extension called GreaseMonkey, works on a wide set of Web systems, but has been extensively tested on Facebook.

INTRODUCTION

Online social networks gained an exponentially growing number of members by offering users new, effective tools for interaction and communication (Isaías 2009). By using Facebook and other social networking systems, people keep in touch each other, meet new friends and connect again with old ones, find jobs and discover new interests. Moreover, users express themselves by sharing textual contents together with pictures, music and videos.

The lack of accessibility of main social networks systems is a widely known issue since these social network services become so pervasive. During the last years just few and partial solutions to overcome this lack were provided by developers (Hailpern 2009, Zajicek 2007). Recently, Facebook accessibility has been enhanced by removing some technological barrier (a CAPTCHA image) from its sign up interface. Currently Facebook accessibility is mainly addressed in a very simple and incomplete way, substantially based on the use of the mobile version of the system. Analogously Twitter, MySpace and other widely used social networks are partially accessible to users with disabilities (AbilityNet 2008) and are not compliant with any national regulation or W3C guidelines (W3C 2008). This accessibility issue is quite common in most of Web 2.0 services, which are strongly based on smart interfaces implemented with a wide use of AJAX scripts (Garrett 2005). Social networks, and more specifically Facebook, support users' customization as the possibility of changing few layout characteristics, such as color background or text size. These modifications are bound to the user profile page only and they have to be manually specified by the user without any support by some profiling tool.

In this paper we present GAP4APE (GreaseMonkey And Profiling for Accessible Pages Enhancement), a system to improve page accessibility based on GreaseMonkey augmented browsing (Greaspot 2011). The system works by automatically modifying the page content on the basis of a profile describing the user's needs and preferences (Mirri et al. 2011). This task is carried out on the basis of a GreaseMonkey through transcoding techniques in order to provide the user an adapted and optimized version of each page, according to his profile. Such modifications are applied both to static content (like text, images or embedded multimedia) and to dynamic content (as AJAX scripts) that controls the interface dynamics. In order to describe user's needs and preferences in GAP4APE, we have used a profiling system, based on the well-known IMS ACCLIP standard (IMS 2002a). GAP4APE can support user in navigating every site by means of a set of customized client-side transcoding scripts that could be developed by a community of users. GAP4APE activates the appropriate sets of script on each page, if a specific group of transcoding activities is available. Otherwise, a standard set of real-time adaptation mechanisms is applied, thereby improving accessibility of all the sites browsed by the user.

We provide a set of scripts devoted to improve accessibility of Facebook interface as a test suite and case study for the effectiveness of the system. In order to design and develop scripts which improve Facebook accessibility, some people with disabilities have been invited to report how they use their assistive technologies/tools while navigating Facebook. This group of users has been involved during the scripts design phase and also during the testing one.

The reminder of this paper is organized as follows. The next Section ("*The GAP4APE System*") presents main design issues and the system architecture. Section entitled "*Facebook accessibility*" describes main accessibility issues of Facebook, while Section entitled "*GAP4APE at work*" shows how our system works on the case study of Facebook. Finally, "*Conclusions and Future Work*" Section ends the paper by introducing main conclusion and future work.

THE GAP4APE SYSTEM

GAP4APE is a client-side augmented browsing system designed and implemented with the aim of enhancing Web pages accessibility by dynamically and automatically updating them to specific user's needs. The design of GAP4APE has been driven by the idea that "*the best Web content for each one*" philosophy should replace the more traditional idea of accessibility claiming that "*one Web content for everyone*" (Salomoni et al. 2008). On one hand, Web accessibility principles assert that using more than a Web page to provide differentiated content to users with disabilities is intended as a discriminating and segregation factor (Bohman 2003). On the other hand, with the pervasive diffusion of client-side technologies, the customization of content has become a widely use technique to improve usability. Moreover, it is worth noting that a single accessible Web page, providing a unique content, could not be optimized for each user who browses it (Salomoni et al. 2008). Concepts and techniques for content transcoding offer the means of generating optimized versions of a same primary resource, providing each user with content adapted to his/her needs and preferences (Bigham 2007, Bigham and Ladner 2007).

User profile

On the basis of the user's profile, GAP4APE performs a set of transcoding scripts. The system provides the user with a browser interface to set his/her preferences and needs defining his profile. To specify a profile that effectively describes special needs of users with disability (including characteristics and settings of assistive tools) we have used a significant part of a well-known standard, the IMS ACCLIP (Accessibility for Learner Information Package) (IMS 2002a).

IMS ACCLIP is a part of the IMS Learner Information Package (IMS LIP) specification (IMS 2002b). In such a part, a set of packages is defined, which can be used to import/export data into/from an IMS compliant e-learning platform. In particular, in IMS ACCLIP the user is described in terms of accessibility needs by the means of a XML-based syntax. Practically speaking, ACCLIP describes user's preferences and needs (visual, aural of device), in order to allow a learning content customization on the basis of them (e.g. preferred/required input/output devices and/or preferred/required content alternatives). Thus, such a profile could be exploited as a description of how users interact with an e-learning environment by focusing on accessibility requirements (Mirri et al. 2011). The ACCLIP specification defines the required elements to represent accessibility preferences, which can be grouped into the following sections:

a. *display information:* this section describes how the user prefers to have information displayed or presented;
b. *control information*, this section defines how a user prefers to control the device;
c. *content information*, this section describes what enhanced, alternative or equivalent content the learner requires;
d. *accommodations*, this section allows recording of requests for and authorization of accessibility accommodations for testing or assessment.

GAP4APE profiling considers attributes from the *a, b* and *c* section, excluding the ones of the *d* section. In particular, our profiling system groups the preferences and needs information into Text, Color, Audio, Visual and General sets.

```
...
<accessForAll schemaVersion="1.0.29"
xmlns="http://www.imsglobal.org/xsd/acclip"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-
instance" xsi:schemaLocation=
"http://www.imsglobal.org/xsd/AccessForAllv1p0.
xsd">
<context identifier="userX" xml:lang="it">
   <display>
     <screenReader>
        <screenReaderGeneric>
          <link value="speakLink"/>
          <link value="differentVoice"/>
          <speechRate value="500"/>
          <pitch value="0.8"/>
          <volume value="0.5"/>
        </screenReaderGeneric>
     </screenReader>
     <braille>…</braille>
   </display>
   <control>
       <keyboardEnhanced> …
       </keyboardEnhanced>
       <mouseEmulation>…</mouseEmulation>
       <voiceRecognition> …
       </voiceRecognition>
   </control>
</context>
</accessForAll>
...
```

Figure 1: Fragment of a profile set by a blind user

Let us consider a blind user who gains access to the Web with a PC equipped with a screen reader and a Braille display (i.e., the assistive technologies that enable blind people to use a computer). A simplified portion of ACCLIP derived elements comprised in the user's profile is reported in Figure 1. In particular, a set of preferences is specified, related to the use of the screen reader (see element `<screenReader>` inside `<display>` element), as well as the Braille display characteristics (see `<braille>` element, partially omitted). All these elements are included inside the accessibility LIP element (`<AccessForAll>`) which drives the system transcoding process. The system adequately transcodes the Web pages on the basis of such profile.

Transcoding

On the basis of the above cited profile, GAP4APE automatically modifies the page content, by adapting it to the chosen features. This activity is performed by a set of transcoding scripts, selected by the system considering both the user profile and the page content. Recently, augmented browsing technologies offer the opportunity to create effective client-side transcoding applications. We have decided to use GreaseMonkey, a browser extension that allows users to develop, install and use scripts which make on-the-fly changes to HTML DOM page (Greaspot 2011, Pilgrim 2005). Since its first release, GreaseMonkey was used to support Web users in automatically adapting Web pages for accessibility purposes. Currently GreaseMonkey is also supported by some mobile devices (Greaspot 2011).

In order to enhance the accessibility of Web content and to provide the best adaptation to each user, by meeting his/her needs and preferences, our scripts allow the transcoding of Web pages by modifying the CSS rules, the HTML DOM and also Web 2.0 scripts. Some examples of such adaptations applied by our scripts will be described in Section *"GAP4APE at work"*.

The system recognizes the Web application browsed by the user and verifies the availability of specialized scripts to perform a page-specific adaptation. If such a set of script is not available, GAP4APE applies a general set of default scripts to improve the page accessibility. Our system adaptation mechanisms works in analogy with well-known screen readers (i.e. Jaws (Freedom Scientific 2011)). These assistive tools act on desktop applications applying customized scripts, whenever they are available and default scripts otherwise. It is worth noting that screen reader users frequently write their own scripts to improve accessibility of applications and share them with other users. Analogously, it is possible for GAP4APE users to write specific adaptation scripts to be shared with other users, as in a crowdsourcing community.

Architecture

The following Figure 2 shows GAP4APE architecture: each user is provided (at the client-side) with a browser operating the GreaseMonkey extension.
GAP4APE works as a set of GreaseMonkey scripts, structured into two modules: *ProM*, Profiling Module, devoted to store and manage the user profile, and *TraM*, content Transcoding Module that performs the adaptation of content.

Once the page from the server has been downloaded, GAP4APE TraM verifies the set of scripts to be applied (a specific one or the more general default scripts set) and matches it with the profile defined by the user. The new version of the page, adapted by considering the user's needs and preferences, is displayed to the user. The profile can be edited and managed through the GAP4APE ProM directly by the user through the browser interface.

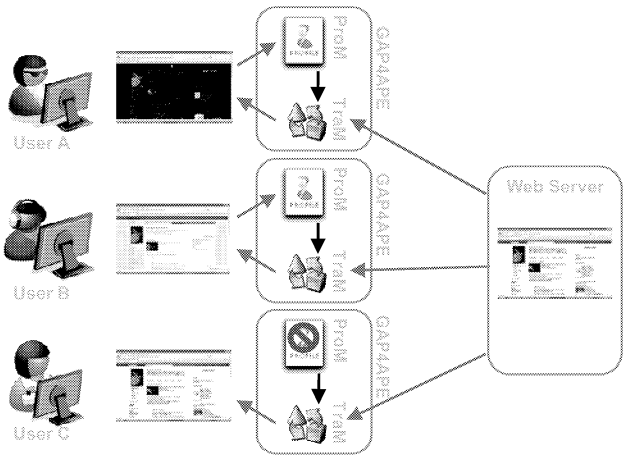


Figure 2: GAP4APE Architecture

Figure 2 shows how three users, once they set their different profiles, browse three different Facebook pages, derived from the same original one. More specifically:

- User A, with low vision, receives a high contrast page, adapted to be browsed by a screen magnifier user.
- User B, who is blind, receives a page with a simplified and linearized structure, adapted to be browsed by a screen reader user.
- User C, who as an empty profile, receives the Facebook page as it is, without any adaptation.

Details on adaptation performed on Facebook page are reported in the following Section *"GAP4APE at work"*.

A GAP4APE prototype has been developed and it has been tested by a group of users with visual impairments (blind people and people with low vision). The entire system will be completed as a set of different extensions tailored for different Web browsers. In this way, GAP4APE will improve Web pages accessibility independently from the browser the user exploits to navigate. The current prototype of GAP4APE has been implemented for Mozilla Firefox, on the basis of the GreaseMonkey mechanism which is devoted to build extensions (Greaspot 2011).

FACEBOOK ACCESSIBILITY

Since its birth Facebook has shown several accessibility problems, which have been mainly due to the heavy use of Web 2.0 and scripting technologies. In particular, a study which was conducted in 2008 on main online social network accessibility (AbilityNet 2008) has rated Facebook *"very inaccessible"* and it has shown that main Facebook accessibility lacks are related to the following issues:

a. Facebook required the user to identify a CAPTCHA image when creating an account. This was a barrier for vision impaired users, particularly screen reader users, but also users with minor vision impairments or dyslexia.
b. There was no separate accessibility page on any of the websites tested. Many websites now display an accessibility page as a way of making a public statement of commitment to ensuring disabled access to their website and to explain any measures,

such as access keys which allow keyboard users quick access to important pages. In addition, an accessibility page can be used to provide extra information for disabled users, such as contact details for the organization in question.

c. Facebook requires JavaScript to be enabled otherwise a user cannot create an account. Scripts can often cause difficulties for those using older browsers, those with vision impairments using some special browsers, and those whose organizations disable JavaScript for security reasons.
d. If an existing Facebook user logs into the website, yet has JavaScript disabled, there will be various functionalities unavailable to them such as drop down menus not working and the ability to set your phone up to send photos directly to your Facebook account.
e. Text size on every page cannot easily be resized by many users – so vital for many visitors who have vision impairment or who are viewing the website on a small screen. A user should override the default text size by changing their browser settings, some website content can overlap and make the text difficult or impossible to read.
f. Users of screen reading software pull all links on a page into a list – so they can quickly access the link they want. Many links do not make sense when read out of context in this way, such as many occurrences of "See all" or "Change" that each takes the user to a different page, or perform a different action.
g. Keyboard only users would experience varying degrees of difficulty, ranging from a lack of links that allow them to jump over main navigation links, to pages or features that were effectively inaccessible to keyboard users.
h. Many graphics lack any alternative text (or "tool tips"). This is true of both graphics that are essential to navigation as well as graphics used purely for cosmetic purposes, causing problems especially for blind screen reader users and that using voice recognition.
i. Although Facebook offered the option to users of linking and including videos, it does not mention the importance of captions (or subtitles), and it does not give the option of adding a transcript to a video.

Even if in the last years Facebook developers have worked to solve these issues, some of the main accessibility problems remain. In particular, the CAPTCHA image is currently shown together with an auditory one, so that users with visual impairments can sign up to the system. Moreover, Facebook added a page to present its main accessibility features. Though this page, Facebook team suggests to users with visual disabilities the use of the mobile HTML-only version. Currently issues *a*, *b* and *e* have been solved.

The previous report (AbilityNet 2008) shows that main accessibility problems on Facebook are related to users with visual disabilities; hence we have concentrated our efforts in defining scripts to provide accessibility solutions for such user.

In order to improve Facebook accessibility, we have analyzed it in its current release and we have identified main problems with the support of users with disabilities. Some solutions have been applied by developing scripts for GAP4APE. Our analysis has begun with the study of how people with disabilities use assistive technologies to enjoy Facebook (Facebook 2011).

A discussion group of people with disabilities has been involved, answering to some interviews and participating in the design. Another group of users has been involved in the testing phases. It is worth noting that each user has got a proper way to navigate the Web and to use his/her own assistive technology. Let us take into account blind users: they subjectively enjoy the features of screen readers, depending on their experiences, their skills, their knowledge about such tools and the frequency they use them.

From a sample of 16 blind users who have been interviewed about how they navigate the Web through a screen reader, we know that all of them use the combination of TAB key and arrows keys, only six of them are used to search text by using the combination of CTRL key and F key, only one of them uses the combination of INS and F6 keys to obtain a list of the headings in the Web page and, finally, 10 of them use the combination of INS and F7 keys to obtain a list of all the links in the Web page.

In general, blind users who navigate Facebook pages through a screen reader face different levels of barriers and meet different problems. Some of these main problems can be summarized as follows:

- The chat is not accessible.
- Headings are not well-organized and their hierarchy is not clear.
- Some links provide a cyclic navigation, without a clear destination.
- Some important and useful features and parts of the content are difficult to be reached.
- Useless information and images makes the navigation difficult and heavy.
- Some text links are ambiguous.
- Some links and some information are redundant.
- Some useful features are read as simple text instead of button titles, links or labels (e.g. the "Comment" feature).
- There are some difficulties in finding friends when coincidences of names happen.
- Each update refreshes the whole page.

A similar set of issues has been identified by studying the interaction with Facebook of a group of users with low vision and users with color blindness.

GAP4APE AT WORK

This section is devoted to describe the Facebook use case. We will illustrate how a blind user and a low vision user can set their preferences and how they can enjoy the social network.

Setting preferences

The user can improve the navigation experience by defining his/her needs and preferences through the Preferences Panel (PP). In the current version of GAP4APE prototype, such a Panel is available as a menu choice in the Firefox interface and it is provided to users as a window which is displayed over the browser one.

The Panel window has been created by using XUL (XML User Interface Language) (Mozilla 2011), which is an XML user interface markup language developed by the Mozilla Project. XUL is a Mozilla's XML-based language that permits to build feature-rich cross platform applications. Such applications can be customized with different text, graphics and layout. XUL applications are based on other W3C standard technologies too, such as HTML 4.0, CSS 1 and 2, DOM Levels 1 and 2, JavaScript 1.5, ECMA-262 Edition 3 (ECMAScript). Plus, XUL takes into account also the W3C eXtensible Bindings Language (XBL), which is a markup language to define special new elements, or "bindings" for XUL widgets. XBL allows developers to extend XUL by tailoring existing tags and creating new ones. Thus, developers can define and create customized user interface widgets. XUL provides a clear separation among the client application and programmatic logic (XUL, XBL and JavaScript), presentational aspects (CSS and images) and language-specific text labels (DTDs). This way, the layout and appearance of XUL applications are independent from the application definition and logic. This enabled us to define a Preferences Panel which is accessible itself.

The PP organizes all the configurable characteristics into the following sets: Text, Color, Audio, Visual and General. Figure 3 and Figure 4 show screenshots taken from the Color and the Visual Tabs of the Preferences Panel.

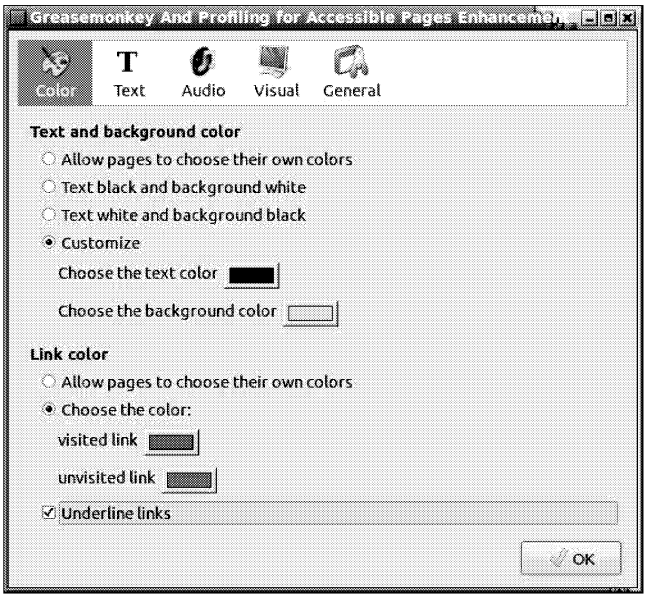


Figure 3: The Preferences Panel: the Color Tab

Through the Preferences Panel, the user can also choose and fix traditional browser preferences. This way, the user can have a more complete view of all the kind of customizations he/she could enjoy and can configure a wide set of characteristics, by using only one tool. Such a mechanism allows users to choose and set just once (for instance, during the browser installation) his/her preferences and needs, so as to enjoy adapted and accessible Web content anytime he/she access the browser. Obviously, through the Preference Panel, the user can change his/her preferences whenever he/she wants. Such new preferences will be immediately available. User's configurations are set through JavaScript and saved by the Preferences Storage System (Mozilla 2011).

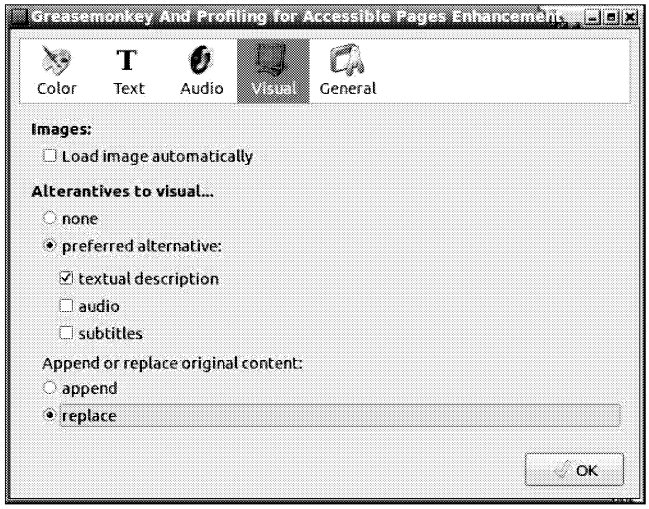


Figure 4: The Preferences Panel: the Visual Tab

Adapting Facebook

The GAP4APE content transcoding system is based on a set of scripts which have been developed by using JavaScript. Such scripts interact with the Web page DOM and they add, remove or change elements, on the basis of users' preferences and needs, as they are declared into the profiles. For instance, it is possible to remove links or images or any other kind of visual elements such as advertisements, to change the links text, to show the alternative descriptions instead of the related images or to add them and show them close to the related images. Moreover, changes to the CSS are applied by using JavaScript too. Some scripts can directly modify single CSS rules, while some others can create a new CSS file, by applying the user's preferences and needs, and then they substitute the old CSS with the new one in the HTML code of the Web page. With this kind of scripts it is possible to modify layout and presentational aspects, such as the colors used in the Web page or the font family. Finally, also script transcoding is realized by using JavaScript. Some of these scripts are developed to automatically modify Web 2.0 scripts (for instance to avoid automatic refreshing and/or updating of the Web page), while some others are designed to be executed only on a proper Web application, since it is not always possible to automatically identify and modify AJAX scripts in a feasible and effective way. Hence our prototype applies suitable scripts to transcode specific Web application pages (in particular when the users request Web 2.0 social networks content and services), in a way which is similar to screen readers behavior with desktop applications.

In order to increase Facebook accessibility we have designed and developed a set of scripts which face the problems

described in section "FACEBOOK ACCESSIBILITY". In particular, our scripts:

- Re-label text links in order to avoid ambiguous links.
- Remove redundant links and information (in particular into Users' profile pages).
- Label form elements.
- Remove useless images from Users' profile pages and from the Wall.
- Provide a more accessible chat.
- Block the automatic updating and allow users to choose when refreshing the page.
- Assign and reorganize the headings hierarchy.
- Reorganize lists and nested list items.
- Reorganize the whole layout of the page, grouping in a fixed area all the advertisements and all the information which makes uneasy the navigation with a screen reader.

When a blind user declares his/her preferences and needs by the means of the GAP4APE Preferences Panel, he/she could choose to substitute images with textual alternatives or the preferred kind of alternative (textual, auditory, etc.). Figure 5 shows a screenshot of the original Facebook wall, while Figure 6 represents the same wall with the application of GAP4APE scripts.

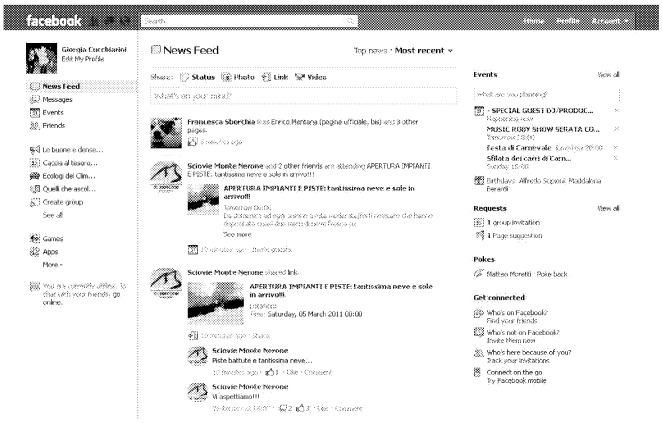


Figure 5: The Facebook wall

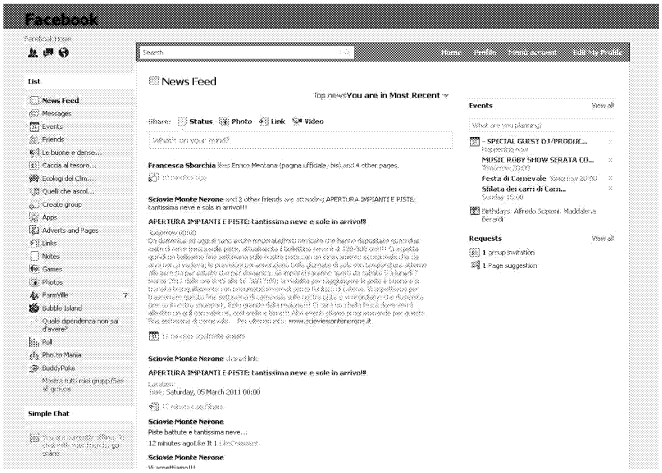


Figure 6: The accessible Facebook wall

Now let us take into account users with low vision. They face different problems; in particular, it is very difficult to meet their needs, since there are several kinds and levels of these visual impairments. The Preferences Panel allows users to set a wide group of settings in a very detailed manner, so as to better transcode the Web pages and to better meet users' needs and preferences.

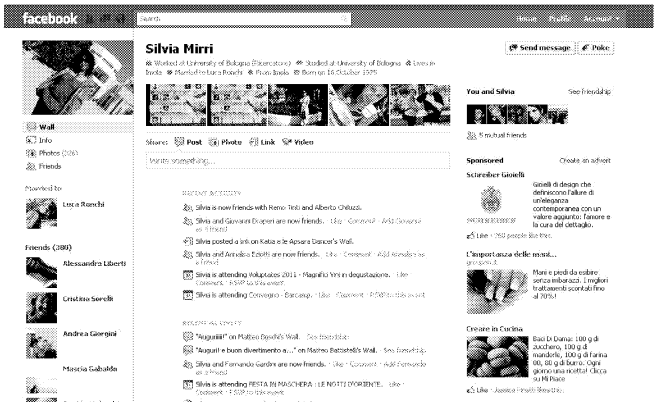


Figure 7: The Facebook user's profile

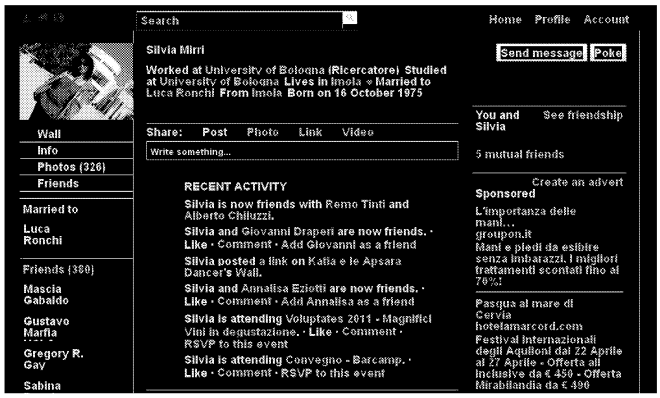


Figure 8: The Facebook user's profile adapted for low vision

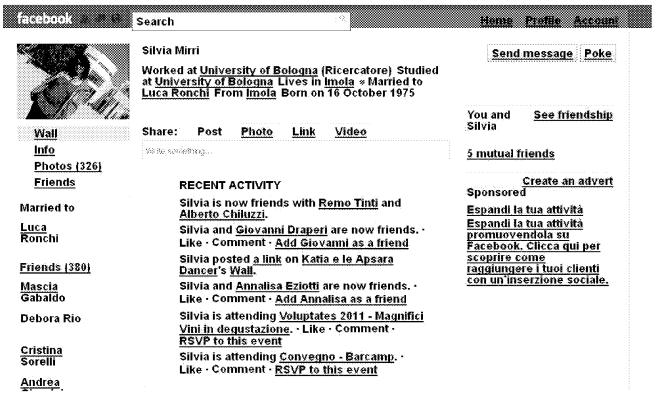


Figure 9: The Facebook user's profile adapted for low vision

An accessible version of Facebook User's Profile page is depicted in Figure 8: through our GAP4APE scripts the user has set different text and background colors with a high contrast and a bigger text size. In fact, the original color contrast is not enough for user with low vision, moreover

black background and yellow texts provides a good level of contrast.
Finally, Figure 9 shows another accessible version of the same Facebook User's Profile depicted in Figure 7 and Figure 8. In this case the applied scripts increase the text size, changing only the color of links (in order to make them more visible) and they reorganize the Web page layout so as to support users with low vision by the means of a simpler and easier to navigate interface.

In order to evaluate our GAP4APE prototype we have conducted a testing phase. We have involved 10 users with visual impairments and asked them to complete a set of task on Facebook pages (such as: to comment a specific post on the wall, to modify profile and privacy settings, to remove a post from the profile page, etc.) by the means of GAP4APE scripts. After that they evaluation testers have answered to a questionnaire. According to the questionnaire answers all the users have completed all the tasks, 70% of testers have found easier to comment and remove posts, while 30% did not found differences between the original version of Facebook and the customized one, 80% of tester prefer transcoded Facebook pages in order to modify profile and privacy configurations and all the users have found easier the whole interface of Facebook wall and profiles pages.

CONCLUSIONS AND FUTURE WORK

In this paper we have presented GAP4APE, a client-side transcoding system with the aim of adapting Web pages in order to enhance Web accessibility, by meeting user's preferences and needs.
To reach this goal, we have faced two different issues:

1. Profiling users' preferences and needs. Dealing users with disabilities we have defined a profiling system which is based on the well-known IMS ACCLIP standard.
2. Transcoding Web pages. We have exploited a two-layers system which applies: (i) a specific set of scripts devoted to a given Web application when such scripts are available; (ii) a default set of scripts otherwise. Transcoding activities are performed on the client-side: a Web page is delivered to any user, but GAP4APE adapts it by transcoding Web content, CSS rules, the HTML DOM and also Web pages scripts. The two-layer transcoding system answers to the need of automatically identifying and modifying AJAX scripts. This kind of adaptation has to be performed by sets of scripts, specifically designed for a given application (see 2.i). Default scripts are not effective in adapting code in a unique and feasible way. The idea of ad hoc scripts for specific applications has been inspired by the behavior of well-known screen readers, such as Jaws.

This paper illustrates a use case of GAP4APE in improving the accessibility of one of the most common and widely used Web 2.0 social network: Facebook. In order to design and implement scripts which enhance Facebook accessibility, it has been analyzed how people with disabilities exploit their assistive technologies while they navigate Facebook. This group of users has been involved during the design phase and also during the testing one.
Future works will be mainly addressed to the integration of our system into a wider set of browsers (including Chrome) and to the definition of Web services which provide automatic content transcoding, involving also multimedia ones, in order to overcome JavaScript limits in providing complex content transformation. Moreover, specific scripts will be designed and developed in order to improve other Social Network applications, such as Twitter and LinkedIn.

ACKNOWLEDGEMENTS

The authors wish to thank Ludovico Antonio Muratori and Catia Prandi for their precious support.

REFERENCES

AbilityNet Web Accessibility Team. 2008. "State of the eNation web accessibility reports: Social Networking Websites." Available from: http://www.abilitynet.org.uk/docs/enation/2008SocialNetworkingSites.pdf

Bigham, J.P. 2007. "AccessMonkey: enabling and sharing end user accessibility improvements". ACM SIGACCESS Accessibility and Computing - ASSETS 2007, No. 89, 3-6.

Bigham, J.P., and Ladner, R.E. 2007. "Accessmonkey: A Collaborative Scripting Framework for Web Users and Developers." In *Proceedings of the 2007 international cross-disciplinary conference on Web accessibility (W4A2007)* (Banff, Alberta, Canada, May 2007). ACM Press, New York, NY, 25-34.

Bohman. P. 2003. "Introduction to Web Accessibility." Available from: http://www.webaim.org/intro/

Facebook. 2011. Available from: http://www.facebook.com/

Freedom Scientific. 2001. "JAWS for windows." Available from: http://www.freedomscientific.com

Garrett, J.J. 2005. "AJAX: A new approach to Web applications." Available from: http://www.adaptivepath.com/ideas/essays/archives/000385.php

Greaspot. 2011. "GreaseMonkey Firefox Extension." Available from: http://Greasemonkey.mozdev.org/

Hailpern, J., Guarino-Reid, L., Boardman, R., Annam, S. 2009. "Web 2.0: blind to an accessible new world." In *Proceedings of the 18th International Conference on World Wide Web (WWW '09)* (Madrid, Spain, April 2009). ACM Press, New York, NY, 821-830.

IMS Global Learning Consortium. 2002a. "IMS Learner Information Package Accessibility for LIP." Available from: http://www.imsglobal.org/specificationdownload.cfm

IMS Global Learning Consortium. 2002b. "IMS Learner Information Profile (LIP)." Available from: http://www.imsglobal.org/specificationdownload.cfm

Isaías, P., Miranda, P., Pífano, S. 2009. "Critical Success Factors for Web 2.0 --- A Reference Framework". In *Proceedings of the 3rd International Conference on Online Communities and Social Computing*: Held as Part of HCI International 2009 (OCSC '09) (San Diego, California, USA, Jul. 2009). Lecture Notes in Computer Science, Springer, Berlin, Germany, 354-363.

Mirri, S., Salomoni, P., Prandi C. 2011a. "Augment browsing and standard profiling for enhancing Web accessibility." In *Proceedings of the 2011 international cross-disciplinary conference on Web accessibility (W4A2011)* (India, March 2011), to appear.

Mirri, S., Salomoni, P., Roccetti, M., Gay, G.R. 2011b. "Beyond Standards: Unleashing Accessibility on a Learning Content Management System." Transactions on Edutainment, Springer, Vol. V, LNCS 6530, 35-49.

Mozilla Developer Network. 2011. "XML User Interface Language (XUL)." Available from: https://developer.mozilla.org/En/XUL

Pilgrim, M. 2005. "GreaseMonkey Hacks: Tips & Tools for Remixing the Web with Firefox." O'Reilly Media.

Salomoni, P., Mirri, S., Ferretti, S., Roccetti, M. 2008. "A multimedia broker to support accessible and mobile learning through learning objects adaptation." *ACM Transactions on Internet Technology (TOIT)*, Vo. 8 No. 2, ACM Press, 9-23.

W3C. 2008. "Web Content Accessibility Guidelines 2.0." http://www.w3.org/TR/WCAG20/

Zajicek, M. 2007. "Web 2.0: hype or happiness?". In *Proceedings of the 2007 International Cross Disciplinary Conference on Web Accessibility (W4A2007)* (Banff, Alberta, Canada, May 2007). ACM Press, New York, NY, 35-39.

AUTHOR BIOGRAPHY

SILVIA MIRRI is an assistant professor of Computer Science at the Department of Computer Science of the University of Bologna. Her research interests include: multimodal interaction, accessibility and e-learning.

PAOLA SALOMONI is an associate professor of Computer Science at the Department of Computer Science of the University of Bologna. Her research interests include the following: distributed multimedia systems, wireless multimedia, e-learning accessibility.

GIORGIA CUCCHIARINI is a computer science undergraduate student. She is working on her degree thesis which concerns Web 2.0 social networks accessibility.

WEB-BASED LEARNING CONTENT MANAGEMENT SYSTEM OF POWER ELECTRONICS

Valery Vodovozov and Zoja Raud
Tallinn University of Technology
Ehitajate tee 5
19086 Tallinn, Estonia
edrive@narod.ru

KEYWORDS

Learning management, E-learning, Object-oriented approach, Curriculum, Thesaurus

ABSTRACT

This paper focuses on a new information system for web-based management of learning Power Electronics. Learning content management architecture is proposed which effectively utilizes methodical resources of higher educational institutions, particularly their curricula, syllabi, schedules, studies, etc. The proposed system helps to prepare basic modules for particular teaching methods and to emphasize the disciplines without doubling. Important results of the management system's implementation are described.

INTRODUCTION

Learning content management systems (LCMS) have become widely accepted in universities and colleges. It is a multi-user environment where the participants of an educational process can create, store, use, manage, and deliver digital learning content from a central object repository. LCMS has the following benefits:

- it imparts the main knowledge of the field based on fundamental concepts
- it involves techniques to apply the basic knowledge into the solution of the current and emerging problems
- it is suitable for searching effective educational ways along with the changing learning situations

This paper describes an effective web-based learning management framework developed for Power Electronics and tools that the authors have used to redesign and implement in the electrical engineering program at Tallinn University of Technology (TUT). The paper describes the main components of the developed LCMS and presents an educational thesaurus which serves as the core of the toolkit proposed.

MANAGEMENT OF LEARNING CONTENT

Learning content management is a related technology to learning management being its further development. In the multi-user LCMS environment the developers can maintain not only training modules but also all the individual pieces that make up learning content. An LCMS is proposed to deal at the "atomic" knowledge level being focused on the content design, control and publishing. LCMS applications allow users to create, import, search for and reuse the small units of digital learning content and assets, commonly referred to as learning objects.

A learning object is a collection of content, practice, and assessment items that are combined based on a single learning objective. Learning objects carry many names, including content objects, chunks, educational objects, information objects, intelligent objects, knowledge bits, knowledge objects, learning components, media objects, reusable curriculum components, nuggets, reusable information objects, reusable objects, and testable units of cognition, training components, and units of learning. Learning objects which consist of the small "granules" of learning are maintained in the frame of databases called repositories.

These assets may include media files developed in other authoring tools, assessment items, simulations, text, graphics or any other object that makes up the content within the course being created. They typically have a number of other components, which range from descriptive data to information about rights and educational level. All learning objects have the following key characteristics:

- being a new way of thinking rather than the traditional content which comes in a several hour chunk, learning objects are much smaller units, typically ranging from 2 to 15 minutes
- being a self-contained unit, each learning object can be taken independently
- being a reusable cell, a learning object may be involved in multiple contexts for different purposes
- being an aggregated unit, learning objects can be grouped into larger collections of content, including traditional course structures
- being tagged with metadata, every learning object has descriptive information allowing it to be easily found by a search

Contemporary LCMS provide tools for authoring and reusing or re-purposing their content as well as the virtual spaces for student interaction, such as discussion forums, live chat rooms, and web-conferences. The studies in (Cox 2010; Agrawal et al. 2008; Lachiver and Tardif 2002; Rasseneur et al. 2003; McShane et al. 2001) indicate that with an appropriate application of the LCMS technology

and matching to a complete instructional plan for design and use of learning objects, significant efficiencies can and will be achieved.

An LCMS manages learning content across the company training areas. It supplies the staff and students by the means to create and use learning content without doubling of efforts. In the remote hosting approach, an LCMS may host the content in a central repository and allow multiple LMS to access it. The primary business problems that an LCMS solves are as follows:

- centralized management of learning content for efficient searching and retrieval
- productivity gains around rapid and condensed development timelines
- outcomes regarding the assembly, maintenance and publishing, branding and delivery of learning content

In an LCMS the focus is on the authoring and management of e-learning reusable content. Rather than developing entire courses for multiple audiences, LCMS provide the ability for single course instances to be modified and republished for various audiences. The objects stored in the centralized repositories commonly are available to course developers and content experts throughout an organization for potential repurpose and reuse, thus allowing for the rapid assembly of customized educational needs

The described LCMS of Power Electronics was implemented at the Department of Electrical Drives and Power Electronics of TUT. The objectives of the LCMS introduction into the course syllabus were to expand significantly the learning opportunities of different groups of students, both the strong and the weak ones. To develop the system, an object-oriented approach has been applied (Vodovozov and Raud 2009).

The core learning objects of the LCMS are stored in the repository of Estonian national e-learning portal available at: *http://www.e-ope.ee/repositoorium*. The system involves the Web-textbook on Power Electronics (Vodovozov 2010), hypertext tutorial aids of exercises and laboratory practice, videos helping to start laboratory works and exercises, current assessment sheets, a list of examination problems, and some other documents.

The learning flash videos “Getting Started Diode Rectifiers”, “Getting Started Thyristor Rectifiers”, and “Getting Started IGBT Converters” describe the basic components and the related rules of laboratory experimentation. As the stands are complicated electronic devices, the videos help to prepare the laboratory works during the work preparation, before the real tests are performed. Implementing experiments with virtual learning objects provides an opportunity to visualize and explore the complex equipment before the lab doing. Another flash video “Getting Started Multisim” serves as an instrument to learn about using the Multisim toolkit in the study of power electronic converters. The video teaches the students how to develop, produce, and simulate power circuits, to detect and remove errors, and to prepare standard reports, thus providing training before the real design of power electronic converters. Here, the students learn to make schematic models in Multisim, select appropriate components, and combine them in the proper order. Later, the students launch the real simulation to understand electrical processes in power electronic converters.

A package of the three learning objects manages the implementation of the laboratory works. The objects “Diode Rectifiers”, “Thyristor rectifiers” and “Transistor DC converters” include detailed guidance, circuit diagrams, formulae, and questions.

A separate group of learning objects is devoted to the simulation procedures in the Multisim toolbox. These learning objects involve onscreen instructions with the hyperlinks to the Power Electronics thesaurus, thus providing modelling, experimentation and reporting. Commonly, the process guided by the learning objects of such kinds consists of the following steps:

- development and calculation of power electronics circuits
- selection of electronic components
- schematic building
- analysis of voltage and current timing diagrams
- explanation and documentation of the results obtained
- assessment of the properly drawn up schemes and their diagnosis

A virtual learning object “Simulation of Single-Phase Half-Wave Rectifiers” offers an opportunity to prepare and develop simulators of the single-phase half-wave rectifiers of five different circuits: simple M1 rectifier, M1 rectifier supplied a resistive load, M1 rectifier with LC filter, M1 rectifier with inductive load, and M1 thyristor rectifier. An object “Simulation of Single-Phase Full-Wave Rectifiers” teaches the students how to develop simulators of the single-phase full-wave rectifiers resulting in seven different circuits: simple M2 rectifier, M2 rectifier with LC filter, M2 thyristor rectifier, simple B2 bridge rectifier, B2 rectifier with LC filter, B2 rectifier feeding an inductive load, and B2 thyristor rectifier. In the object “Simulation of Three-Phase Rectifiers”, instructions are given concerning the implementation of the rectifier simulators that include four three-phase circuits: simple M3 rectifier, M3 thyristor rectifier, simple B6 rectifier, and B6 thyristor rectifier. An object “Simulation of AC Converters” covers the preparation and making of four AC converters: a single-phase voltage regulator, a single-phase bridge inverter, voltage source and current source inverters, and a frequency converter. The last object of this group “Simulation of DC Converters” instructs virtual dc/dc converter design in seven different varieties: step-down chopper, closed-loop buck chopper, step-up chopper, closed-loop boost chopper, combined step-down and step-up chopper, closed-loop buck-boost chopper, and Cuk chopper.

A special self-assessment module was prepared as a combination of Web pages and Excel sheets to provide student's self-assessment in Power Electronics. The object

includes the rating table accompanied by the student's self-evaluation rules. As the learning process involves the mandatory tasks and the optional problems that add scores into the learner's rating, this object consists of three important parts: the results of the on-site lecture tests, laboratory works and exercises. During the semester, the students may follow online their current rating and their expected examination grade. Using these data, they obtain the tool to plan, adjust, and predict learning outcomes. Particularly, by solving additional tasks, they can improve the personal rating and the final grade.

THESAURUS OF POWER ELECTRONICS

All electronic documents are accompanied by an educational thesaurus of Power Electronics (Raud and Vodovozov 2009). Using interactive hyperlinks, this object clarifies and explains the sense of learning objects through other learning materials, such as lectures and practical guidelines.

The thesaurus of Power Electronics is a hierarchically structured interactive dictionary that interprets more than 700 terms and concepts in the field of Power Electronics. Each of its records has a semantic (meaningful) relationship with the preliminary given definitions. The thesaurus is represented using the database structure and the Concept Map toolbox.

The frame of the thesaurus includes the following areas:

- concept name
- concept definition
- "More..." area with a hyperlink to the appropriate part of the electronic textbook with the detailed explanation of the concept and required references
- "Feedback" area where the predetermined concepts are listed arranged in the form of hyperlinks with the appropriate thesaurus entries
- "Feedforward" area where the inheritable concepts are listed arranged in the form of hyperlinks with the appropriate thesaurus entries

An alphabetically ordered index of the thesaurus is settled in the starting records of the thesaurus. Another choice is a thematic index that may be used to guide the learner through the course from the root concept to the leaves of the knowledge tree.

The concept map (Novak and Canas, 2008) serves as the graphical organizing tool of the thesaurus. It includes concepts and relationships between the concepts. The concepts are represented in a hierarchical fashion. The hierarchical structure for a particular domain of knowledge also depends on the context in which that knowledge is being applied or considered. The concept map of the Power Electronics thesaurus comprises six files interconnected by the hyperlinks: "Introduction", "Rectifiers", "Inverters", "AC/AC converters", "DC/DC converters", and "Control circuits".

The thesaurus enables the learners to overcome the barrier between the practical application and the theoretical knowledge. It promotes learning and ultimately students' progress and achievement and has a major influence on what learners learn, how effectively they learn, and consequently on the quality of their learning.

CONCLUSION

An effective learning content management framework based on a total perception of curricula and syllabi development of Power Electronics is described. The novel toolkit was used by the authors to redesign and improve the Power Electronics program at TUT. The main components and learning objects of the proposed LCMS are cited in the Internet. The powerful educational thesaurus serves as the core object of the educational system.

ACKNOWLEDGEMENT

This paper is supported by Estonian Science Foundation Grant ETF8020.

REFERENCES

Agrawal, S., Sharma, P. B. and Kumar, M. 2008. "Knowledge management framework for improving curriculum development processes in technical education", *3rd International Conference on Convergence and Hybrid Information Technology ICCIT'08*, Busan, Korea, vol. 2, pp. 885-890

Cox, D. 2010. *USA Government Funding Options for Corporate Training: Understanding the Landscape... Knowing Your Options.* Brandon Hall Research Associate and CEO, USA.

Lachiver, G. and Tardif, J. 2002. "Fostering and managing curriculum change and innovation", *32nd Annual Frontiers in Education Conference FIE'02*, Boston, MA, vol. 2, pp. F2F-7-F2F-12

McShane, E.A., Trivedi, M. and Shenai, K. 2001. "An improved approach to application-specific power electronics education: Curriculum development", *IEEE Transactions on Education*, vol. 44, no. 3, pp. 282-288

Novak, J. D. and Canas, A. J. 2008. "The theory underlying concept maps and how to construct and use them", *Technical Report IHMC CmapTools.* Florida Institute for Human and Machine Cognition, USA.

Rasseneur, D., Jacoboni, P. and Tchounikine, P. 2003. "Enhancing a web-based distance-learning curriculum with dedicated tools", *IEEE/WIC International Conference on Web Intelligence WI'03*, Beijing, China, pp. 285-291

Raud, Z. and Vodovozov, V. 2009, "Professional thesaurus of engineering educational system", *2nd International Multi-Conference on Engineering and Technological Innovation IMETI'09*, Orlando, FL, 2009, pp. 212−217.

Vodovozov, V. 2010. *Introduction to Power Electronics*, Ventus Publishing, Denmark.

Vodovozov, V. and Raud, Z. 2009. "An object-oriented approach to curriculum design", *7th International Conference on Education and Information Systems, Technologies and Applications EISTA'09*, Orlando, FL, pp. 225−230.

DATA TRANSFER ARCHITECTURES

PIDS FOR CONGESTION CONTROL IN INTERNET USING OPNET

Teresa Alvarez, Virginia Alvarez, Lourdes Nicolas and Anuar Salim

Dpt. of Systems Engineering and Automatic Control,
Industrial Engineering School (Doctor Mergelina)
University of Valladolid (SPAIN)
E-mail: tere@autom.uva.es, viralvcue@gmail.com, lour.nico@gmail.com, anuarsalimdelrio@gmail.com

KEYWORDS
Simulation, computer networks, congestion control, OPNET, TCP, AQM.

ABSTRACT

This paper presents some simulation-based examples used at graduate level, so that engineering students understand the problems caused by congestion in computer networks, and some possible solutions. This is a necessity, as fast and reliable computer links are needed, but there are situations when congestion happens in those links, caused by fluctuations in the network delay and the number of users. The methodology presented here allows students to learn the importance of the Quality of Service (QoS) and how it depends on the methods used to solve the congestion problem. This work presents a combined approach of ready and self-programmed methods combined with tutorials using OPNET and MATLAB. Different approaches for dealing with congestion are considered: Drop Tail, RED and PIDs.

1 INTRODUCTION

Future Computer Science and Telecommunication Engineers (in Spain) learn about Communications Networks (Internet, TCP/AQM), Finite State Machines (FSM) and automatic control techniques. It is a difficult task to find systems that appeal them simultaneously from the communications and control perspectives.

The congestion control problem in Internet fulfils these requirements: huge amounts of data are transferred from one point to another in a matter of seconds, or even less, but problems are frequent (Azuma et al. 2005). Congestion is responsible for many of these communication inconveniences.

What is it congestion? When does it happen? Congestion occurs when there are too many sources sending too much data too fast for the network to handle, and it is a serious problem. Thus, it is necessary to reduce this problem as much as possible. In (Jacobson 1988) and (Ryu et al. 2004) there are good descriptions of the problem. Feedback control techniques have already been applied to solve the congestion control problem, to improve the Quality of Service (QoS) that the users observe.

Internet congestion control is carried out in the transport layer, whereas the Active Queue Management scheme (AQM) is implemented in the routers. Some AQM algorithms are drop tail, RED and control theory based approaches such as PID (Hollot et al. 2002, Athuraliya et al. 2001, Misra et al. 2000, Srikant, 2004).

In the learning program presented here simulation is central for understanding what happens in reality and comparing all sort of scenarios. In fact, most of the related works in the literature use simulation software, as this makes it possible to compare different techniques is repeatable conditions (Pawlikowski et al. 2002). Apart from using MATLAB/SIMULINK, the students also work with OPNET. Alternatively, other software (EcosimPro, Modelica, Dymola, ns-2, ns-3, etc) could be used.

Why do we work with MATLAB and OPNET as simulation languages? The use of MATLAB/SIMULINK is widespread in Engineering, so the students are familiar with it. The decision between ns and OPNET is not so straightforward: although both simulation tools have advantages and disadvantages, OPNET is more visual, and easier to learn from scratch, so it is the one selected. In summary: MATLAB is used by the students for studying the congestion control problem from the automatic control perspective: transfer functions and step responses. Then, PID controllers can be tuned and compared with classical strategies such as drop tail or RED. Once this is done, the students can progress to deal with more realistic situations: OPNET provides a very realistic environment for testing and simulating networks. As it is based on using FSM, network events are handled as close to reality as possible. Moreover, the possibilities for adding new algorithms are great but not straightforward (the tutorials that the students can download are useful for this).

From this point onwards, students (and in general, every user) can simulate more complex topologies and program new strategies without having to spend quite a lot of time in programming issues. The examples described in this paper are based on a dumbbell topology (Figure 1).

The learning methodology proposed is based on two successive steps that are later described in detail:

1. Using a dynamical model for a congested router, the students derive the corresponding transfer function, which is then used to tune PID controllers. The model provided for the students is the one in (Misra et al. 2000), as it is the most frequently used in practice.

2. The students, using some basic examples proposed, learn to simulate in OPNET a given network topology (such as the one in Figure 1) with the designed PID as congestion control algorithm.

Supplementary handouts and tutorials (available in www.isa.cie.uva.es/~tere/opnet.html) give suggestions for the students on what to change, what to include and where. Particular attention is given to the understanding of these two techniques: RED and PID.

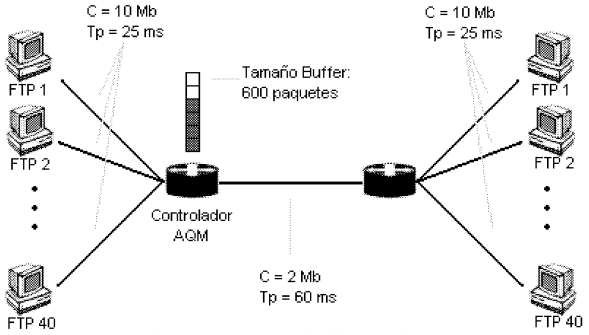


Figure 1: Dumbbell topology

The rest of the paper is organized as follows. Section 2 describes the basics of the first step: the linear model proposed for the students and the Simulink implementation. Section 3 deals with the use of OPNET. Finally, some conclusions are discussed.

2 FIRST STEP: DESIGN USING A LINEAR MODEL

Although an AQM router is a non-linear system, in order to understand the dynamics involved, and design simple controllers, it is proposed to use the linearized model in (Hollot et al. 2002). Using equation (1) (where, N is the number of TCP sessions, C is the capacity of the link in packets/sec. and R_0 is the round-trip time) as model of a congested router and following the classical closed-loop block diagram in Figure 2, the students implement Simulink blocks equivalent to those depicted in Figure 3.

$$P(s) = e^{-sR_0} P_{TCP}(s) \cdot P_{queue}(s) = e^{-sR_0} \frac{R_0 C^2 / 2N^2}{s + \frac{2N}{R_0^2}} \cdot \frac{N/R_0}{s + \frac{1}{R_0}} \quad (1)$$

Initially, the PID controller follows the classic expression given in equation 2, although the student should derive a proper implementation, as described in section 3:

$$K(s) = K_p E(s)\left(1 + \frac{1}{T_i s} + sT_d\right) \quad (2)$$

The control signal is the probability of marking a packet (p) and the error is evaluated from the difference between the desired bottleneck queue length and the real one (q).

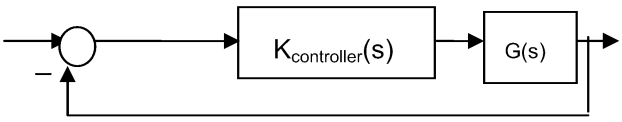


Figure 2: Block diagram

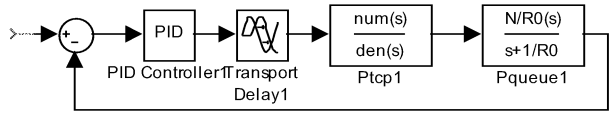


Figure 3: Simulink Blocks

3 SECOND STEP: OPNET

As it has been mentioned, apart from understanding congestion control techniques, OPNET Modeler is used by students, as it is a very powerful network simulator, used in practice for research or development. Communication networks, protocols, and all sort of applications can be implemented and studied with great flexibility. One of the main advantages of OPNET is that it works with FSM, so the students can implement realistic situations with great reliability.

RED (Floyd and Jacobson 1993) is included in the OPNET's distribution package, but not PID controllers. Adding new algorithms requires quite a good knowledge of the structure of OPNET: Implementing the PID controller in OPNET requires good knowledge of the internal configuration of the software, so the details are provided to the students. How can PIDs be implemented in OPNET? Equation (2) represents the classic or textbook formulation, but as the controller is implemented in a computer, some practical issues should be taken into account. Following (Aström and Hägglund 2006), the discrete equations implemented in OPNET are given by (3).

$$\left.\begin{aligned} p(t_k) &= K_p e(t_k) \\ d(t_k) &= \frac{T_d}{T_d + N_{PID} \cdot T_s} \\ &\quad \left(d(t_{k-1}) - K_p N_{PID}\left(y(t_k) - y(t_{k-1})\right)\right) \\ u(t_k) &= sat\left(p(t_k) + d(t_k) + i(t_k)\right) \\ i(t_{k+1}) &= i(t_k) + \frac{K_p}{T_i} e(t_k) \end{aligned}\right\} \quad (3)$$

where N_{PID} is a value between 8 and 20 and T_s is the sampling period.

The PID's source code is not complex, but finding the right place where it should be included in the simulation tool is not so straightforward. Several changes in external

files (such as oms_qm.h, oms_qm.ex.c) have to be made, and also the process model needs modifications. The more challenging modification is in the FSM. The PID controller works at fixed times (sampling time) and not every time a new packet arrives to the router as when using RED or drop tail techniques. This fact has to be taken this into account within the FSM. Thus, the FSM has to be expanded with new states. Moreover the Quality of Service (QoS) node (shown in Figure 4) needed some adjustments. This way, the students can select the parameters of the new AQM algorithm.

PID Status	PID Enabled
Sample Time	0.08
Qref	50
Kp	-0.05
Ki	0.0
Kd	0.0
Type PID	B
Zn	Enabled
ti	0.009
td	0.01
b	1.0
N	8.0

Figure 4: QoS node

Figure 5 shows a possible implementation of the dumbbell topology in OPNET, whereas Figures 6 and 7 depict simulation results for one of the different situations under study: the students should be able to generate and analyze this kind of results. A detailed tutorial with the models and software modifications is available on the web.

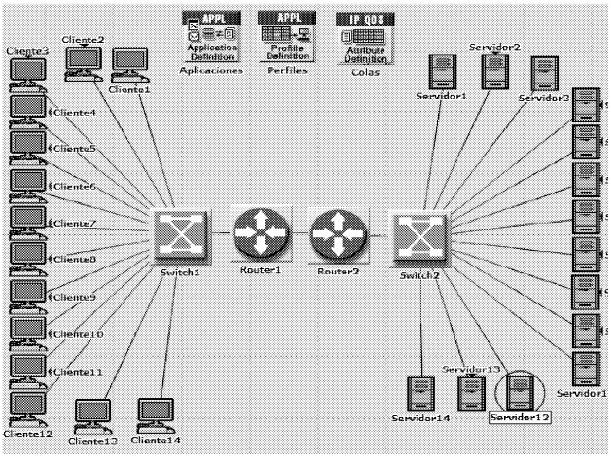


Figure 5: Dumbbell topology in OPNET

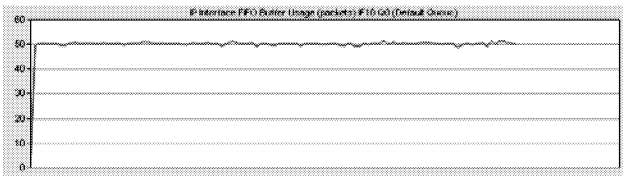


Figure 6: Simulation results (PID)

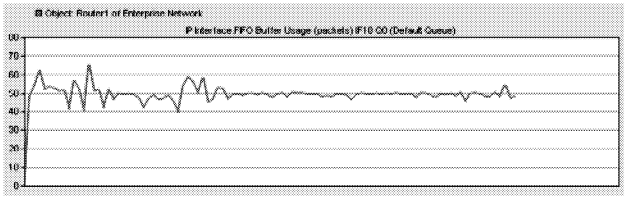


Figure 7: Simulation results (RED)

4 CONCLUSIONS

This paper has presented an approach on how students learn alternative methods for improving the QoS when the congestion problem appears in a network, based on using congestion control algorithms, based on control techniques such as PIDs. MATLAB and OPNET are the software tools chosen to work with.

REFERENCES

Aström, K. J. and T. Hägglund. 2006. "Advanced PID control". ISA. NC.

Athuraliya, S., S.H. Low, V.H. Li and Q. Yin. 2001." REM: active queue management". *IEEE Networking*, *15*(3), 48-53.

Azuma, T., T. Fujita and M. Fujita. 2005. "Congestion control for TCP/AQM networks using State Predictive Control". *Transactions of the Institute of Electrical Engineers of Japan. C*, *125*(9), 1491-1496.

Floyd S. and V. Jacobson. 1993. "Random early detection gateways for congestion avoidance". *IEEE/ACM Transactions on Networking*, *1*(4), 397-413.

Hollot, C.V., V. Misra, D. Towsley and W. Gong. 2001. "Analysis and Design of Controllers for AQM Routers Supporting TCP flows". *IEEE Transactions on Automatic Control*, *47*, 945-959.

Jacobson, V. 1988, Congestion avoidance and control. *ACM SIGCOMM'88* (Stanford, CA, USA), 314-329.

Misra, V., W.B. Gong and D. Towsley. 2000. "Fluid Based Analysis of a network of AQM routers supporting TCP flows with an application to RED". *ACM/SIGCOMM*, Stockholm, Sweden, 151-160.

Pawlikowski, K., H.D. Jeong and J.S. Lee. 2002. "On credibility of Simulation Studies of Telecommunication Networks". IEEE Communications magazine, 40(1), 132-139.

Ryu, S., C. Rump and C. Qiao. 2004. "Advances in Active Queue Management (AQM) based TCP congestion control". *Telecommunication Systems*, *25*(3-4), 317-351.

Srikant R. 2004. "The Mathematics of Internet Congestion Control". Birkhauser Verlag, Boston.

WEB REFERENCES

DYMOLA. http://www.3ds.com/products/catia/portfolio/dymola

MATLAB. Mathworks. www.mathworks.com.

MODELICA. https://www.modelica.org/

Ns-2. http://www.isi.edu/nsnam/ns/

Ns-3. http://www.nsnam.org/

OPNET. www.opnet.com.

A MULTICAST STREAMING ARCHITECTURE BASED ON A BALANCED PEER-TO-PEER TREE OVERLAY

Cristian Petruţ Ene, Eliana-Dina Tîrşa, Mugurel Ionuţ Andreica, Nicolae Ţăpuş
Department of Computer Science
Politehnica University of Bucharest
Splaiul Independenţei 313, sector 6, 060042, Bucharest
Romania
E-mail: cristiene85@yahoo.com, eliana.tirsa@cs.pub.ro, mugurel.andreica@cs.pub.ro, nicolae.tapus@cs.pub.ro

KEYWORDS
Multicast, Streaming, P2P, Tree, Decentralized, Balanced.

ABSTRACT

In this paper we present a novel multicast streaming architecture based on a peer-to-peer overlay with a balanced tree structure. The overlay is maintained in a fully decentralized manner. The system has been fully implemented and evaluated under real network conditions. The results clearly show the improvements brought by our system over a classical client-server architecture.

INTRODUCTION

Multicast streaming is an important data transfer mode, in which a stream of data needs to reach multiple destinations (clients). In the classical client-server model, the server provides the data stream and the clients connect individually to the server in order to receive the stream. This way, if N clients connect to the server simultaneously, the server will have to serve the stream N times in parallel. Peer-to-peer overlays can be employed in order to relieve the stress on the server and let the clients deliver the stream further to other clients (and not simply consume the stream).
In this paper we propose such a peer-to-peer system for application-level multicast streaming based on a tree overlay. All the nodes interested in receiving a stream join the overlay and, besides consuming the stream, they may also forward the stream to their overlay neighbors. In the overlay, every peer may dynamically become a content generator (i.e. publish its own stream) which may be received by the other peers. The tree overlay is maintained as balanced as possible (i.e. its diameter is minimized) and each peer has an upper limit on the number of overlay neighbors it may have.
The rest of this paper is structured as follows. First, we present the architecture of the system and the main functions provided by each peer. Next, we discuss implementation details. Then, we present experimental results. Another contribution of this paper, of a more theoretical nature, is concerned with batch computations of aggregates, which may be performed in a multicast streaming system – we present this contribution closer to the end of the paper. Finally, we discuss related work, conclude and present future work.

SYSTEM ARCHITECTURE

The peer-to-peer overlay is organized as a tree. This structure facilitates multicast data transmission. Every peer can be a content producer. When a peer generates content, this content is sent through the overlay to all the other peers (whenever a peer receives a data packet, it forwards it to all of its neighbors, except for the one from where the packet was received). The tree structure ensures that no unnecessary packets are transmitted. By also imposing an upper limit on the number of tree neighbors of each peer we are able to limit the amount of bandwidth consumed by a peer, thus balancing the bandwidth consumption throughout the overlay.
Each peer has a unique self-generated identifier (the id is generated by using a hash function on a sufficiently large number of bits) and considering several "nearly" unique attributes (e.g. time moment when the peer was started, etc.).

Gossiping

The overlay is maintained through gossiping. Periodically, peers send advertisements to all of their neighbors and to their neighbors' neighbors (i.e. the *extended neighbors*), to let them know that they are still alive (and within the system). This way, each peer will also know its extended neighbors.
Each peer P periodically (re)computes for each neighbor V a set of values. We will denote by V_P the set of neighbors of the peer P and by $S(V,P)$ the subtree of the overlay containing the node P obtained by removing the edge (P,V).
We will denote by $MAX\_DEGREE(P)$ the maximum number of neighbors peer P may have (it is possible for each peer to have a different maximum degree, though we will mainly consider that all peers have the same upper bound) and by $DEGREE(P)$ the current number of neighbors of P. If $DEGREE(P)<MAX\_DEGREE(P)$ then we will say that P has at least one *empty position* (where a new node X may attach to the tree by connecting to the peer P).
We denote by $minDist(V,P)$=the minimum distance in the overlay from V to an empty position in $S(V,P)$. Moreover, we denote by $maxDist(V,P)$=the maximum distance from P to a leaf in $S(V,P)$. We have:

$$\text{minDist}(V,P)=\begin{cases}1, \text{if MAX\_DEGREE}(P) > DEGREE(P) \\ 1+\text{MIN}\{\text{minDist}(P,X) \mid X \in V_P, X \neq V\}, \\ \text{otherwise}\end{cases} \quad (1)$$

$$\text{maxDist}(V,P)=\begin{cases}1, \text{if DEGREE}(P)=1\\ 1+\text{MAX}\{\text{maxDist}(P,X)\mid X\in V_p, X\neq V\}, \\ \text{otherwise}\end{cases} \quad (2)$$

These values are (re)computed at each gossiping period. While gossiping, besides the existence advertisement, each peer *X* sends to every neighbor *P* the values *minDist(P,X)* and *maxDist(P,X)* computed by the peer *X*. Then, at the next gossiping round, every peer *P* will recompute all the values *minDist(V,P)* and *maxDist(V,P)* for each of its neighbors *V*.
The computation of these values is based on the results from (Andreica et al. 2009), where a balanced multicast tree construction method was proposed. In the absence of new peers joining the system and old peers leaving the system, the values *minDist(*,*)* and *maxDist(*,*)* converge to their correct values after a number of gossiping rounds which is proportional to the diameter of the tree. The convergence proof is based on similar arguments to the proof presented in (Andreica et al. 2009) and, thus, we will not detail it here.

Joining the System

When a new peer joins the system, it will enter as a leaf in the tree overlay. The new peer must first know one other peer which is already inside the tree. Then, the new peer will contact the known peer, which will redirect the new peer step by step towards some peer in the tree, which will become the new peer's only neighbor.
The redirection of a new peer is performed by using the values *minDist(*,*)* and *maxDist(*,*)*. After the new peer *P* contacts an existing peer *X*, the peer *X* will redirect *P* to one of its neighbors *Y* for which *maxDist(X,Y)>maxDist(Y,X)*. Then, *P* will contact *Y* and *Y* will proceed the same as *X* did previously. In this stage, *P* is redirected towards a peer which is close to the center of the tree.
Eventually, *P* will contact a peer *Z* which has no neighbor *Y* such that *maxDist(Z,Y)>maxDist(Y,Z)*. Then, the second stage of the joining process begins. If *DEGREE(Z)< MAX_DEGREE(Z)* then *P* will connect directly to *Z*. Otherwise, *Z* will select some neighbor *Y* for which *minDist(Z,Y)* is minimum (among all of *Z*'s neighbors) and *P* will be redirected to *Y*. From there on, *P* will be redirected using the *minDist(*,*)* values (from the contacted peer to one of its neighbors according to the rule mentioned above), until *P* reaches a peer *W* with *DEGREE(W)<MAX_DEGREE(W)*. Finally, *P* will connect to *W* in the tree.

Leaving the System

When a peer *X* leaves the system, it doesn't have to announce anyone. Its neighbors will detect its departure because no more advertisements will be received from *X*. Then, all of *X*'s neighbors contact each other (as they are extended neighbors of each other). Each neighbor *Y* sends to all the other neighbors *Z* the value *maxDist(X,Y)*. The neighbor with the largest sent value is selected as their representative (ties are broken deterministically, based on the peer identifiers). Then, the representative *Y* will contact the leaf which is furthest away from it in *S(X,Y)*. This can be achieved by repeatedly traversing the peers in *S(X,Y)* (a similar procedure is discussed in (Andreica et al. 2009)). Then, the selected leaf *L* will disconnect from the tree and reconnect in the place of the departed peer *X* (i.e. it will connect to all of *X*'s former neighbors).

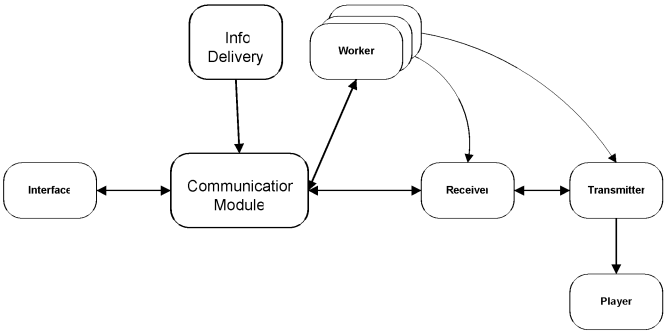


Figure 1: Software architecture of a peer's program.

IMPLEMENTATION DETAILS

We implemented our system using the Java programming language. The code executed by each peer consists of several software modules, presented in Fig. 1.
The communication module is responsible for the data transfer between a peer and its neighbors. The peers communicate through sockets (either TCP or UDP – this is configurable). We used the *java.nio* package (Hitchens 2002) for socket communication. Message writing and message processing is handled by a pool of worker threads. Thus, when another module sends a message, this message is handled to a worker which will write it on the corresponding socket. Moreover, when a new message is read from a socket, it is passed to a worker in order to process it.
The communication module also stores the list of neighbors, together with their network addresses (and, if TCP is used, a list of open TCP sockets to the neighbors). A list of all the data streams (generated locally or received from other peers) is also stored by the communication module.
The main task of a worker is the processing of the received messages and the sending of messages. However, it is also used for handling other tasks, such as:

- establishing a connection to a specified destination
- creating and transmitting a data stream's messages to all the neighbors of the peer ; the source of the stream may be a file or a live capture (capturing live data from a camera was performed using the Java Media Framework API (Gordon and Talley 1999))
- retransmitting the messages of a data stream to all the peer's neighbors except the one from where the message was received
- handling the departure of a neighbor: when a neighbor is detected to have left, it will be removed from the data structures maintained by the peer and the extended neighbors (i.e. the neighbors of the departed peer) will be contacted in order to decide together a new peer which will replace the departed peer

The *Interface* module is used for transmitting user commands to the communication module and for allowing the user to select which of the existing streams to visualize.
The *InfoDelivery* module is responsible for gossiping, i.e. for periodically sending the needed information to the neighbors and advertisements (including the network address and identifier of the peer) to the neighbors and the extended

neighbors.
The *Transmitter* and *Receiver* sub-modules are responsible strictly for the stream data delivery and reception.
The *Player* module plays the received streams which were selected by the user.

EXPERIMENTAL EVALUATION

Join and Leave Tests

We performed several validation tests in the following scenario. At first, 5 peers on 5 different machines were started. The maximum degree of a peer was set to 3. These 5 peers formed the overlay in Fig. 2.

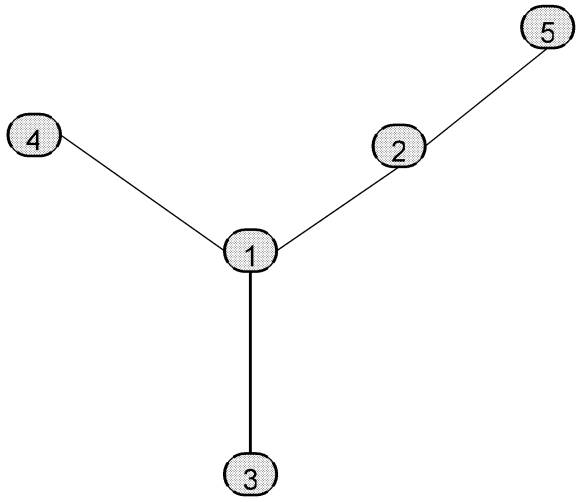


Figure 2: Initial testing scenario.

Then, peer 6 was inserted into the system. Peer 6 knows only peer 3 initially. Peer 6 is redirected as shown in Fig. 3.

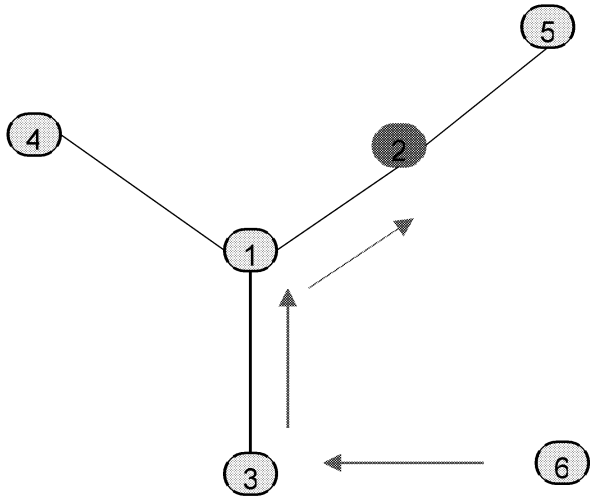


Figure 3: Peer redirection when a new peer joins the tree.

Then, peer 1 is removed from the system. Peer 1's neighbors detect its absence and communicate with each other in order to select a representative. Peer 5 is chosen as the replacement peer for peer 1. Peer 5 will disconnect from its current neighbors and take the place of peer 1 (as shown in Fig. 4).

Comparison to a client-server architecture

A live stream was published by a node and the other 5 nodes all wanted to see the live stream simultaneously. We computed the bit rate and frames per second of each of the 5 clients when the initial node acts as a server and when all the 6 nodes are interconnected in the multicast tree. Table 1 presents the results.

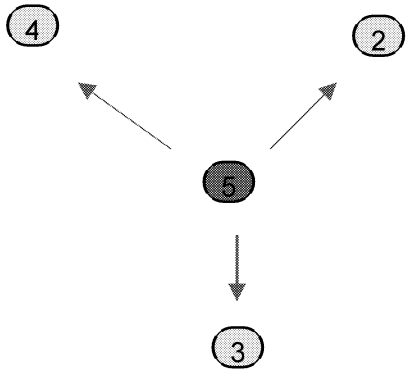


Figure 4: Peer 5 replaces the departed peer 1.

Table 1: Comparison between the proposed system and a client-server architecture

Peer id	Our P2P system		Client-server architecture	
	Bit rate (kbps)	Fps	Bit rate (kbps)	Fps
1	662.481	25	96.92	0.5
2	556.392	24	300.012	13
3	639.338	26	115.72	4
4	619.18	24.5	117.996	1
5	583.71	25	247.513	8

Comparison to an unbalanced P2P tree overlay

Next, we considered the case when peers simply joined the tree at random positions (i.e. they still join the tree as leaves, but their initial neighbor is chosen randomly, instead of according to our rules).

We inserted 20 peers and we set the maximum peer degree to 3. Our system obtained a diameter of 6 (which is the minimum possible for 20 peers), while the random version obtained a diameter of 10. The structure of our system is presented in Fig. 5. Tree diameter has a direct influence on the playback delay between the stream generated by the source and the stream received at the peers – the larger the diameter, the larger the "lag" between the stream's position at the source and its position at the other peers will be.

BATCH COMPUTATIONS OF AGGREGATES

We consider a number N, an initial value $val(1)$ and a non-invertible function f (such that $f(x)$ can be computed in $O(1)$ time for any value x). The values $val(2 \le i \le N)$ are defined as: $val(i)=f(val(i-1))$. We define the values $A(i)$ $(1 \le i \le N)$ where $A(i)$ is an aggregate of all the values $val(i)$ $(1 \le i \le N;\ i \ne N)$. the aggregate function $aggf$ is a commutative but non-invertible function. We want to compute the N values $A(*)$ efficiently.

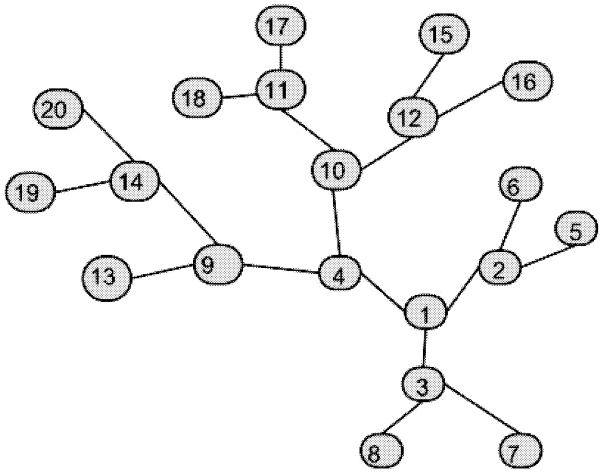


Figure 5: Multicast tree overlay with 20 peers and maximum degree set to 3.

If *aggf* were invertible, an easy solution is the following. First, we compute *aggn* as *aggf(val(1), …, val(N))*. Then, we generate the values *val(i)* again and, for each *i*, we have $A(i)=aggn\ aggf^{-1}\ val(i)$. If the values *A(*)* need to be stored, then this solution uses *O(N)* memory. If they do not need to be stored, then they can be output as soon as they are computed and this solution uses only *O(1)* memory. The time complexity is *O(N)*.

Since *aggf* is not invertible, the most straightforward solution is to generate and store all the values *val(i)* ($1 \le i \le N$) and then compute each value *A(i)* independently. This approach takes *O(N)* memory space and $O(N^2)$ time.

A better solution with *O(N)* memory space is the following. After computing and storing all the values *val(i)* ($1 \le i \le N$), we will compute two arrays: *Aleft(i)=aggf(val(1), …, val(i))* and *Aright(i)=aggf(val(i+1), …, val(N))*. We have *Aleft(0)=Aright(N+1)*=the neutral element of the aggregation function. Then, *Aleft(*$1 \le i \le N$*)=aggf(Aleft(i-1), val(i))* and *Aright(*$1 \le i \le N$*)=aggf(val(i), Aright(i+1))*. Then, we have *A(i)=aggf(Aleft(i-1), Aright(i+1))*. This solution takes *O(N)* time. Some of the mentioned arrays can be computed one element at a time, but the solution still needs *O(N)* memory space.

A space-optimal solution is based on the divide-and-conquer paradigm. Let's consider a function *DC(i, j, val(i), outF)*, where *outF* is the aggregate of all the values outside of the interval of positions *[i,j]*. The initial function call will be *DC(1, N, val(1), neutral element of aggf)*. The function *DC* works as follows. If *i=j* then we have *A(i)=outF*. Otherwise, let *mid=(i+j) div 2* (integer division). We will compute the aggregate *aggimid* of all the values *val(q)* with $i \le q \le mid$ (by generating these values starting from *val(i)*). We will also compute *val(mid+1)* at the end of these iterations. Then, we will call *DC(mid+1, j, val(mid+1), aggf(aggimid, outF))*. After returning from this call, we will compute *aggmidj*=the aggregate of all the values *val(mid+1), …., val(j)*. Then, we will call *DC(i, mid, val(i), aggf(outF, aggmidj))*. Notice how the values *A(i)* are obtained in decreasing order of *i*. This solution uses *O(log(N))* space, because it uses *O(1)* space on each of the *O(log(N))* levels of the recursion stack. The overall time complexity is *O(N·log(N))*.

A final solution with a linear time complexity (i.e. *O(N)*) and *O(K+N/K)* memory usage (where *K* is a function of *N*). is the following one. For simplicity, we will assume that *N* is a multiple of *K* (we can extend the array of values by values equal to the neutral element of the aggregation function). We will generate the values one by one and store the values in an array *v*. Whenever we reach a position *i* which is a multiple of *K*, we compute *Ragg(i)=aggf(v(1), …, v(K))* and then we clear the array *v*. Then, for every position *i* which is a multiple of *K* (in descending order, starting from *i=N-1*), we set *Ragg(i)=aggf(Ragg(i), Ragg(i+K))*. So far, we used only *O(K+N/K)* memory and *O(N)* time.

Then we regenerate all the values starting from the first and add them to the array *v*. Whenever the current position *i* is a multiple of *K* we compute *aggk*=the aggregate of the values *v(1), …, v(K)*, and then *Lagg(i)=aggf(Lagg(i-K), aggk)* (we consider *Lagg(0)*=the neutral element of the *aggf* function). Then, we will compute the values *pvagg(*$1 \le i \le K$*)= aggf(pvagg(i-1), v(i))* (where we consider *pvagg(0)*=the neutral element of the *aggf* function) and *svagg(*$1 \le i < K$*)= aggf(svagg(i+1), v(i))* (where we consider *svagg(K+1)*=the neutral element of the *aggf* function). After this we can compute the values *A(i-K+1), …, A(i)*, as follows: *A(i-K+j)=aggf(Lagg(i-K), pvagg(j-1), svagg(j+1), Ragg(i+K))* (we consider *Ragg(pos)*=the neutral element of the *aggf* function if *pos*>*N*). Then, we clear the array *v*. As can be noticed, the time complexity of this solution is *O(N)* and its memory space is *O(K+N/K)*. Choosing *K=sqrt(N)* makes the memory usage equal to *O(sqrt(N))* (by *sqrt(N)* we denote the square root of *N*).

The problem discussed in this section is mostly of theoretical interest and is related to the encoding of the data of a multicast stream.

RELATED WORK

JXTA (Gong 2001) is a Java framework for developing peer-to-peer applications. It contains all the basic functions required by such applications. However, we chose not to use JXTA for our system's implementation, for two reasons: (1) we did not need all the features JXTA provided ; (2) we needed full control over the overlay construction.

Multicast streaming architectures based on peer-to-peer overlays have been the object of many previous research papers. Based on their overlay structure, they are mostly of two types: (1) with a tree overlay ; (2) with a mesh overlay. Considering another classification criterion, multicast systems with a tree overlay can be classified according to whether they use only one tree or multiple trees.

Single Tree Overlays

ZIGZAG (Tran et al. 2003) constructs and maintains a multicast tree whose diameter is $O(\log_K(N))$, where *N* is the total number of peers and *K* is their maximum degree. Scribe (Rowstron et al. 2001) constructs a multicast tree embedded into a DHT overlay.

Multiple Tree Overlays

Multiple tree overlays, like Splitstream (Castro et al. 2003), CoopNet (Padmanabhan and Sripanidkulchai 2002) and ChunkySpread (Venkataraman et al. 2006) balance the data forwarding roles of the peers. In a single tree overlay, for instance, a leaf node only consumes content and does not

forward it any further. When multiple trees are used, the chance of a peer being a leaf in each tree is very low and, thus, it contributes its resources more to the whole system. CoopNet uses MDC (multiple description coding), which encodes a media stream into multiple sub-streams. Splitstream and Chunkyspread also use such load balancing mechanisms. Compared to single tree overlays, these systems are more resilient to node failures and departures.

Mesh Overlays

Many P2P streaming platforms, like CoolStreaming (Xie et al. 2007), AnySee (Liao et al. 2006), PRIME (Magharei and Rejaie 2007) or DagStream (Liang and Nahrstedt 2006), construct and maintain a mesh overlay. These overlays are somewhat similar to a multiple tree overlay, however, unlike them, a data packet does not have a fixed (static) route in mesh overlays; instead, a packet is routed dynamically within the overlay. Moreover, periodically, peers may exchange data packets and may ask for packets proactively.

Hybrid Tree-Mesh Overlays

mTreeBone (Wang et al. 2007) uses a hybrid tree-mesh overlay, where a tree overlay is used as the backbone, on top of which a mesh architecture is added. The backbone contains stable peers and the mesh overlay contains the unstable peers (thus, churn affects mainly just the mesh overlay, thus being handled effectively).

CONCLUSIONS AND FUTURE WORK

In this paper we introduced a multicast streaming architecture based on a peer-to-peer tree overlay. Moreover, the tree is maintained as balanced as possible by making local decisions only (i.e. in a fully decentralized manner). Experimental results showed that the proposed system behaves significantly better than a client-server system.

Our system can be employed successfully by Internet TV companies in order to deliver TV channels to the clients with reduced bandwidth costs (by making use of the clients' upload bandwidth) – in this case, the clients would have to install a dedicated application on their machines. The system could also be employed in order to facilitate synchronization, data dissemination and quick notifications between multiple servers of a content delivery network, between multiple offices of the same business company or even between different companies. Our system can be used wherever there is a stream of information which needs to be disseminated to multiple participants (for instance, trading prices at various places).

As future work, we consider the possibility of dynamically reconfiguring the tree overlay. For instance, if a leaf peer were able to detect other peers containing empty positions which are closer to the center of the tree, the leaf could try to occupy that position, by asking for permission from the peer owning the empty position. If permission is granted, then the leaf peer will disconnect from the tree and will reconnect on the empty position.

The paper also discussed a theoretical problem concerned with batch computations of aggregates of the values of a sequence given implicitly. A space-optimal $O(N \cdot log(N))$ time algorithm and a time-optimal $O(sqrt(N))$ memory solution were proposed for the problem.

ACKNOWLEDGEMENT

The work presented in this paper has been partially supported by the Sectoral Operational Programme Human Resources Development 2007-2013 of the Romanian Ministry of Labour, Family and Social Protection through the Financial Agreements POSDRU/89/1.5/S/62557 and POSDRU/6/1.5/S/16, and by CNCSIS-UEFISCDI under research grant PD_240/2010 (contract no. 33/28.07.2010), PN II – RU program.

REFERENCES

Andreica M. I.; E.-D. Tîrşa; and N. Ţăpuş. 2009. "Data Distribution Optimization using Offline Algorithms and a Peer-to-Peer Small Diameter Tree Architecture with Bounded Node Degrees". In *Proceedings of the 17th International Conference on Control Systems and Computer Science (CSCS),* vol. 2, 445-452.

Castro M.; P. Druschel; A-M. Kermarrec; A. Nandi; A. Rowstron; and A. Singh. 2003. "Splitstream: High-Bandwidth Multicast in a Cooperative Environment". In *Proceedings of the International Symposium on Operating Systems Principles (SOSP).*

Gong, L. 2001. "JXTA: A Network Programming Environment". *IEEE Internet Computing*, vol. 5 (3), 88-95.

Gordon R. and S. Talley. 1999. *Essential JMF: Java Media Framework*. Prentice Hall, Upper Saddle River, NJ.

Hitchens R. 2002. *Java NIO, First Edition*. O'Reilly.

Liang J. and K. Nahrstedt. 2006. "DagStream: Locality Aware and Failure Resilient Peer-to-Peer Streaming". In *Proceedings of the Multimedia Computing and Networking Conference*, 224-238.

Liao X.; H. Jin; Y. Liu; L. M. Ni; and D. Deng. 2006. "AnySee: Peer-to-peer Live Streaming". In *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM).*

Magharei N. and R. Rejaie. 2007. "PRIME: Peer-to-Peer Receiver-Driven Mesh-based Streaming". In *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM).*

Padmanabhan V. N. and K. Sripanidkulchai. 2002. "The Case for Cooperative Networking". In *Lecture Notes in Computer Science*, vol. 2429, 178-190.

Rowstron A.; A. M. Kermarrec; M. Castro; and P. Druschel. 2001. "Scribe: The Design of a Large Scale Event Notification Infrastructure". In *Proceedings of the 3rd Inernational Workshop on Networked Group Communication*, 30-43.

Tran D. A.; K. A. Hua; and T. Do. 2003. "ZIGZAG: An Efficient Peer-to-Peer Scheme for Media Streaming". In *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, 1283-1292.

Venkataraman V.; K. Yoshida; and P. Francis. 2006. "Chunkyspread: Multi-Tree Unstructured Peer-to-Peer Multicast". In *Proceedings of the IEEE International Conference on Network Protocols*, 2–11.

Wang; F.; Y. Xiong; J. Liu. 2007. "mTreebone: A Hybrid Tree/Mesh Overlay for Application-Layer Live Video Multicast". In *Proceedings of the International Conference on Distributed Computing Systems (ICDCS).*

Xie S.; B. Li; G.Y. Keung; X. Zhang. 2007. "Coolstreaming: Design, Theory, and Practice". *IEEE Transactions on Multimedia*, vol. 9 (8), 1661-1671.

FUTURE BUSINESS MANAGEMENT IN COMMERCIAL PRODUCTS OF ROAMING SERVICE

Hossam EL-DDIN
University Assis. Professor
Candidate: Scientist Year STU 2008, Bratislava, SLOVAKIA
Alexandria, EGYPT
E-mail: Dr.EL-DDIN@hotmail.com

KEY WORDS
Mobile Access Router, Parameter Identification, Roaming Service, EPA Modeling, Petri-Nets, Numerical Simulations.

ABSTRACT

The Cisco Series Mobile Access Router (MAR) allows entire networks to roam; and it delivers always on IP connectivity for networks in motion. However, the traditional default value of expiration lifeTime in configuring such routers, set to a fixed integer of 100 s while the priority set to 110 s, restricts the system functioning. The main idea of this paper is to address the management of timing factors of waiting and expiration affecting the network performance, in the name of end-to-end packet delay, by: modeling-supplement the modified Mobile-IP protocol based CIA strategy using equilibrium point analysis technique, analyzing it with the aid of Petri-Nets and further reporting the numerical-simulation results, using MathCAD Professional, from a method for minimizing expiration lifeTime factor. The novel aspects in this paper are to adaptively configure the expiration lifeTime within a range less than the traditional default, with a future maximum of only 89.0000002 s, which is much lower than the default, based on the current position of mobile node in the home network, and with directions on practical development, as new advantage updates in configuring commercial products of Cisco MAR.

INTRODUCTION

The transport infrastructure of mobile Internet is moving towards a model of high-speed (Bing 2000) mobile routers; interconnected by wireless core networks. The architectural choices for interaction between IP network layers of these routers, specifically, the time-signaling aspects in communication, are maturing (Bing 2000), (Raab et al. 2005). At the same time, a consensus has emerged in the industry on utilizing Mobile-IP (MIP)-based protocols (Jamalipour 2003) for roaming (Bing 2000) control plane.

MIP is an open standard, defined for wireless communications by Charles PERKINS in the Internet Engineering Task Force (IETF) RFC 3220 (Perkins 2002). By using MIP, mobile node (MN) can keep the same IP address, stay connected, and maintain ongoing applications while roaming between IP networks. MIP is scalable for the Internet because it is based on IP—any media that can support IP can support MIP (Raab et al. 2005).

A MN that can roam while appearing to a user to be at its home network (HN) can be a mobile–host or–router (Dip 2006). Examples of mobile host (MH) include: the laptop computer— that can change its point of attachment from one network or subnet to another. This MN can travel from link to link and maintain communications using the same IP address. There is no need for any changes to applications, because the solution is at the network layer rather than the physical layer (Jamalipour 2003); this provides the transparent communication in network mobility to the MAR.

Some Cisco MARs (Cisco Systems Inc. 2009), (Raab et al. 2005) deliver *always on* IP connectivity for networks in motion. These routers are intended to be mounted on a vehicle. They support Cisco IOS Mobile Networks (IOS "Internetworking operating system"), and allows them to hide the IP roaming from the local IP nodes. This allows IP hosts on a mobile network to connect transparently to the parent network while a router is in motion. For example, an airplane equipped with Cisco Series MAR can fly around the world while passengers stay connected to the Internet.

Since recent trends in wireless and mobile networks (Ozcan et. al. 2010) show that loss or delay of data in the network mobility is critical to system functioning, this paper addresses the timing factors (TFs) affecting the network performance (i.e., delay of earlier work (EL-DDIN 2007)) by: modeling-supplement the modified MIP protocol based CIA "Communication Inter Agents" using equilibrium point analysis (EPA) technique (Chan 2000), analyzing it with the aid of Petri-Nets (ARTIS Software Corporation 2001) and further reporting the numerical-simulation results, using MathCAD Professional, from a method of minimizing expiration TFs. Rather than configuring expiration lifeTime (T_E) to a fixed integer default of 100 s while priority is set to 110 s (Cisco Systems Inc. 2009), novel aspects of an adaptive range with various values, less than traditional default, are presented in the semi-final section of this paper.

RELATED EARLIER WORK

In (EL-DDIN 2007), (Jamalipour 2003), the need for designing an adaptive modeling prototype for tracking a roaming laptop spreads over a wide range of network parameters: network dimensions, propagation time, lifeTime, traffic signaling and management techniques, …. etc.

By identifying each parameter of roaming losses and devising a method (Fisher and MEIER-HELLSTERN 1992), (Chan 2000) to minimize it, the unacceptably high delays, and in sequence losses, associated with real-time traffic over MIP should be able to be brought into the acceptable range.

MIP-CIA Modeling Overview

As shown in Figure 1, for MIP-CIA protocol, each agent in the mobile Internet is modeled as a single server partial queue, with two independent classes of arrivals:

1. Messaging of registration requests (RRs) for initiating the roaming service of a MH from it's HN to a foreign network (FN), with no queuing, is Class-1.
2. Messaging of data packets due to the Internet connection or forwarding service of traffic to the MH, with a suffer from queuing according to the Markov-modulated Poisson process (MMPP) (Chan 2000), (Fisher and MEIER-HELLSTERN 1992), (Philippart 2003), is Class-2.

Modeling assumptions assume that Class-1 has a priority over Class-2 to increase the possibility of initiating a full-duplex connection setup between home agents (HAs) and foreign agents (FAs).

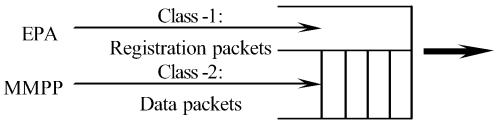


Figure 1: Modeling the Agent (A)

The system is modeled and analyzed with the aid of EPA as a unified and powerful technique for dynamic behavior evaluation of complex packet broadcast networks in different states, for performance evaluation. In EPA, it is assumed that the system is always at an equilibrium point of the Markov chain of network states. In this case, the EPA principle can be applied for the *EtE-D* as in Equation (1):

$$EtE\text{-}D = 2\cdot(\eta + T_P) + T_W + \gamma\cdot\sum_{K=1}^{T} K\cdot\left[1-\gamma\right]^{K-1} \quad (1)$$

Where	
γ	: Probability of receiving/sending an ACK.
T_P	: Propagation time delay between nodes (in s/Km).
T_W	: Waiting/Wasting time between agents (in s/Km).
T_E	: Expiration lifeTime of RR (in s/Km).
η	: Fetching time of data from the MBCT (in s).
K	: Geometric distribution factor.
$EtE\text{-}D$	: End-to-end packet delay (in s/Km).

This work next evaluates the average *EtE-D*, *Av. EtE-D*, for approximated T_E in the design calculations of MIB-CIA.

BASIC PERFORMANCE SIMULATION RESULTS

In more abstracted way, MIP-CIA represents lower *EtE-Ds* (see Table 1 and Figure 2).

Table 1: Comparison of *EtE* delays in ms for MIP-CIA against MIP ones.

	HS		Between HS & FS-1		FS-1		Between FS-1 & FS-2		FS-2	
	From	To	From	To	From	To	From	To	From	To
MIP	35	100	NC	NC	70	130	NC	NC	62	118
MIP-CIA	0.001	0.035	0.035	0.042	0.042	0.065	0.065	0.072	0.072	0.09

Results show that the max. *EtE-D* with CIA procedure is much smaller (better) than the min. *EtE-D* without CIA one.

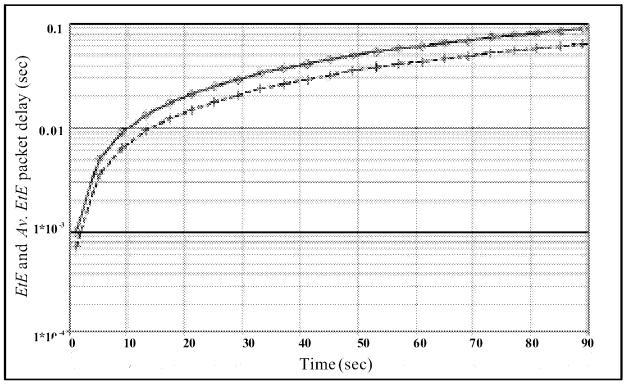


Figure 2: End-to-End packet Delay "*EtE-D*" (up) and Average End-to-End packet Delay "*Av. EtE-D*" (down) versus Waiting Time "T_W".

A=100 agents, *Unit*=1 μs, η =0.01 unit, T_P=0.1 unit/km, for approximated integers of T_E>($2T_P$+T_W).

MOBILE ACCESS ROUTER

Figure 3 shows the three components of Cisco Mobile Networks (MAR, HA, and FA) and their relationships (Cisco Systems Inc. 2009), (Raab et al. 2005) within the mobile network. The MAR functions similarly to the MH with one key difference—the MAR allows entire networks to roam. These MNs can be without MIP client software. The MAR eliminates the need for a MIP client. The MAR "hides" the IP roaming from the local IP nodes so that the local nodes appear to be directly attached to the HN.

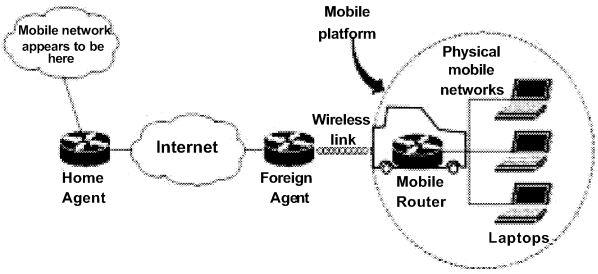


Figure 3: Cisco Mobile Network Components

MAR functionality includes:

- Configuration of the HA address and registration parameters, as well as the interfaces used for roaming.
- Re-registers before the lifeTime expires.

When the MAR is brought up, it is in unknown state. After it receives agent advertisements, it determines if it is at its home or away from its home. It is at home when the source of the advertisement is on its HN, and it behaves like a normal router. When the MAR is roaming, it does not send any routing updates on its roaming interfaces. It sends a RR through the learned FA, to its HA.

The registration lifeTime specifies the duration of a registration for which MAR can attempt to register.

Key Terminologies

- *The registration Lifetime* command specifies the requested lifeTime of each registration. The smallest value between the configured lifeTime and the FA advertised registration

lifeTime is used. The default is 65534 to ensure that the advertised lifeTime is used, excluding infinite. The range is between 3 and 65535 seconds (which represents infinite).

• *The registration Retransmit* command determines how to respond to retransmissions when no reply received and the MAR is not registered with the FA. The initial parameter specifies how long to wait the first time before retrying when no reply is received. The default is 1000 milliseconds (1 second). The range is 10 to 10000 milliseconds (10 seconds).

• *The expire Parameter* is number of seconds to send RR before expiration. The default is 120 seconds. The range is 1 to 3600. If no reply is received, MAR sends another RR after the interval expires. The default is 10 seconds.

FACTORS AFFECTING PERFORMANCE

Since loss or delay of data is critical to system functioning (Dip 2006), (Ozcan et. al. 2010), three main time factors affecting the system delay, as in Equation (1) above, come to the fore and shown in Figure 4:

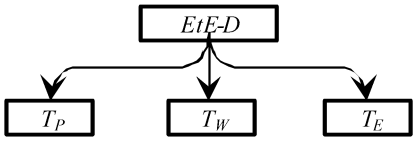


Figure 4: Time Factors affecting Performance

1. *Communication Time / Propagation Time* (T_P): Communication time in wireless mobile networks is controlled by the T_P in the media. T_P delay is the time required for a signal to travel from one point to another on this media. Ad hoc and also mobile networks (Ozcan et. al. 2010), which are based on wireless links between entities of communication, are subject to many new aspects that are generally not observed in infrastructure based networks. For wireless networks, a signal traveling 500 meters and the next time 1 km, will have a larger T_P of about 1.7 μs. For wireless mobile server-client in (Hodge 1996), the links were 2000 km long and had a T_P of 3.7 μs per km. In wireless domain signal T_P is very low. For small ad-hoc network it can be as low as 100 ns (Bing 2000), (Chodorek and Glowacz 2002).

2. *Expiration Mechanism* (T_E): The expiration lifeTime is a number of seconds (in time units) to send the message-alert RRs before expiration, where $T_E>(2T_P+T_W)$ is integer and T_W is the waiting time between agents, until the MN reaches the FA coverage area, where $T_W>=T_P$, ideally if $T_W=T_P$ (EL-DDIN 2007). The traditional default is 120 s and the range is 1 to 3600 s for Cisco platform. For MIB-CIA, he registration items were configured to set the maximum registration lifeTime value of 90 s.

UPDATES METHOD (MODELING SUPPLEMENT)

In this section we present the main contribution of the paper. As mentioned previously, modeling the MIP-CIA based EPA formulates the registration lifeTimes to $T_E>(2T_P+T_W)$, which actually restricts the network performance to *approximated* values during the analysis-design for *Av. EtE-Ds*.
In the corresponding transition states of Petri-Nets (ARTIS Software Corporation 2001) for an EPA-based modeling-supplement, shown in Figures 5, the model describes the system in discrete events. An agent can be in any stateplace and remains in that stateplace for a geometrically distributed amount of time, based on the token length (where token lengths are used to model connection holding times of circuits), and transitions implement the activities of the system, whereas stateplaces are data warehouses that store information until some transition will need it. Before entering the online state *OLS* and while the MN is still within it's HN in the way to the FN and going through the alert state *AS*, from the HA to the chosen FA, the effect of registration parameters T_E and T_W comes to the fore.

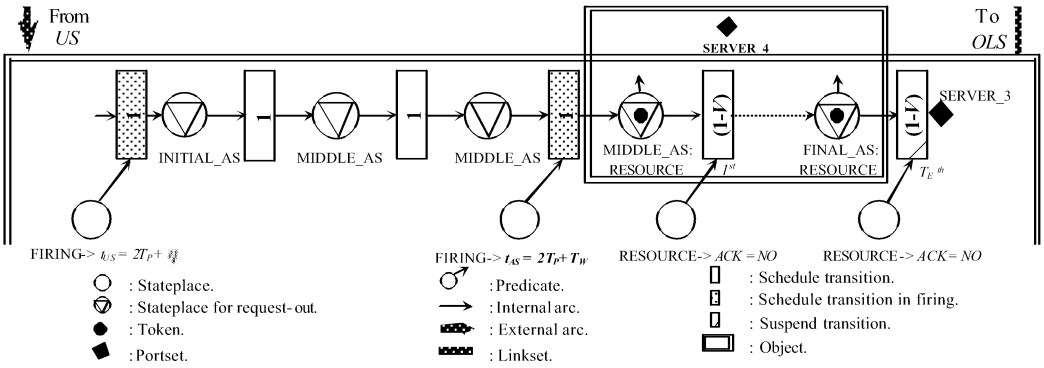


Figure 5: State-Transitions corresponding Petri-Nets for AS

Management method to provide new updates is to minimize the values of T_E. Consult Figure 4, for threshold fractions of T_E, the following Equation (2) holds for exact results:

$$T_E=(2T_P+T_W) \quad (2)$$

UPDATES SIMULATION RESULTS

For different traveling distances *d*, i.e., different T_P, Table 2 proceeds to new updates for T_E and Table 3 reports the new exact results for Equation (2) above.

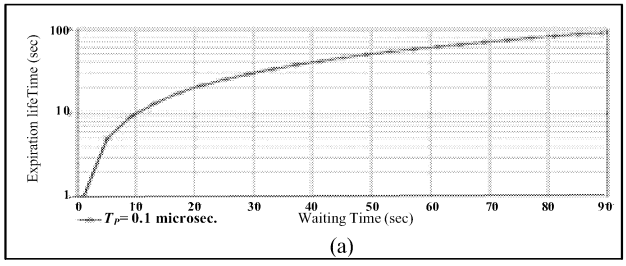
Table 2: Propagation Times for different Distances

	Traveling Distance (Network Dimension) "*d*"		
	< 1km	1 km	2000 km
Propagation Time "T_P"	0.1 μs	1.7 μs	3.7 μs
Expiration LifeTime "T_E"	T_{E1}	T_{E2}	T_{E3}

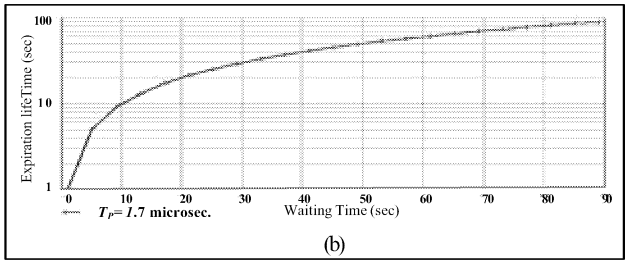
Table 3: Waiting Times "T_W" and corresponding Expiration lifeTimes "T_E" for different Traveling Distances "d". T_W= 1,5, …, 89 sec.

--Waiting Time-- T_W=	----------Expiration lifeTimes---------- $T_{E1}(T_W)$ =	$T_{E2}(T_W)$ =	$T_{E3}(T_W)$ =
1	1.0000002	1.0000034	1.0000074
5	5.0000002	5.0000034	5.0000074
9	9.0000002	9.0000034	9.0000074
13	13.0000002	13.0000034	13.0000074
17	17.0000002	17.0000034	17.0000074
21	21.0000002	21.0000034	21.0000074
25	25.0000002	25.0000034	25.0000074
29	29.0000002	29.0000034	29.0000074
33	33.0000002	33.0000034	33.0000074
37	37.0000002	37.0000034	37.0000074
41	41.0000002	41.0000034	41.0000074
45	45.0000002	45.0000034	45.0000074
49	49.0000002	49.0000034	49.0000074
53	53.0000002	53.0000034	53.0000074
57	57.0000002	57.0000034	57.0000074
61	61.0000002	61.0000034	61.0000074
65	65.0000002	65.0000034	65.0000074
69	69.0000002	69.0000034	69.0000074
73	73.0000002	73.0000034	73.0000074
77	77.0000002	77.0000034	77.0000074
81	81.0000002	81.0000034	81.0000074
85	85.0000002	85.0000034	85.0000074
89	89.0000002	89.0000034	89.0000074

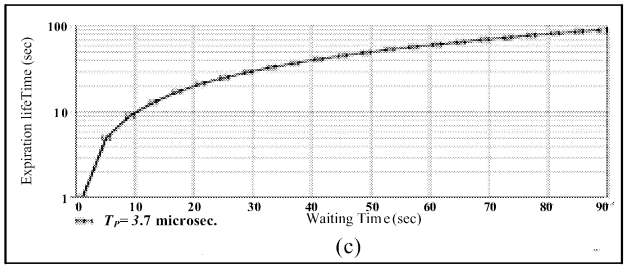
Figures 6-(a) to -(d), according to the T_P of the access media for different ds, with the aid of MathCAD Professional simulations illustrate the numerical adaptive increase in T_E with increasing the T_W, on a time scale, and this behaviour satisfies each MN according to it's current position in the HN before reaching the coverage area of FN, as MIP-CIA based protocol states.



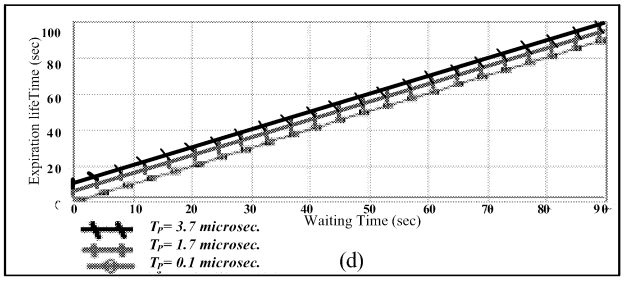
(a)



(b)



(c)



(d)

Figure 6: Expiration lifeTimes "T_E" versus Waiting Times "T_W", for different Propagation Times "T_P" and Exact Fractions of T_E=(2Tp+Tw).

Expiration lifeTime ranges between 1 and exactly 89.0000002 s, with a maximum value of only 89.0000002 s, which is much lower than the default of 100 s while the priority is set to 110 s.

EXPERIMENTAL UPDATES

In practical aspects, our work proceeds toward practical implementation in research labs with the following heuristic of initial O/P configuration file that is to be created during the setup session, with underlined parameters to be updated (Cisco Systems Inc. 2009), (Raab et al. 2005), on an experimental test-bed of Cisco platform: Router-1760/Dram-64 MB/Flash-32 MB using IOS 12.3 (15)T, TFTP server S/W with an access point Cisco Aironet-1230G and a client card with series Cisco Aironet-350G, as in Figure 7.

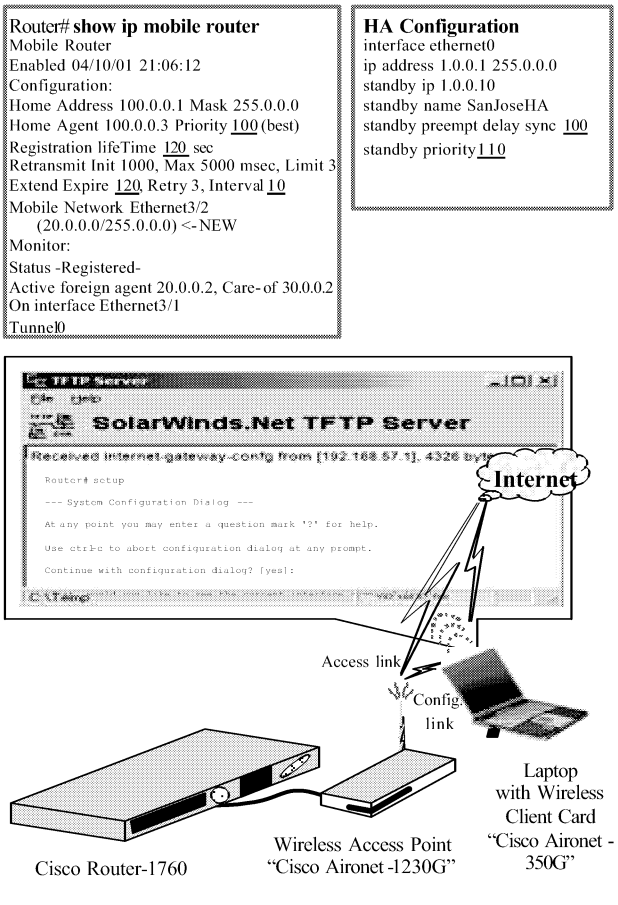


Figure 7: Experimental Config. Test-bed on Cisco platform

Downloading updated S/W image is done in ROM Monitor (ROMMON) mode by using TFTP (**tftpdnld** command). We use command–line interface (CLI) to set the parameters for Cisco IOS S/W loaded on the router. Configuration modes in CLI allow us to make changes to the running configuration and the parameters that govern behavior of routers.

CONCLUSION AND FUTURE WORK

With the actuality that earlier work, using equilibrium point analysis technique (EPA), achieved that registration lifeTime (Expiration parameter) for MIP based CIA performs better than basic MIP, since it ranges between 1 and 90 s, with a maximum value of only 90 s, which is much lower than the default of 120 and range of 1 up to 3600 for Cisco platform. However in this paper, dependency of such lifeTime on other timing factors, that are the propagation- and waiting-times, in defining the chosen configuration value for the expiration lifeTime (EIT) came to the fore. A modeling-based EPA supplement, management method, Petri-Nets analysis and numerical-simulations using MathCAD Professional of such dependence in affecting the network performance were reported, analyzed and proceeded to update-results, that: Adaptively, allowed increased EIT validates the increasing of waiting times and this behaviour satisfies each MN according to it's current position in the HN, ranges between 1 and exactly 89.0000002 s, with a future max. of only 89.0000002 s, which is much lower than the default of 100 s while the priority is set to 110 s; and with directions on practical development. These results show new advantage updates in configuring Cisco mobile access router (MAR).

Ongoing work for the near future concentrates on improving the quality of roaming service by updating the internal kernel of routers configuration, using the developed test-bed. The numerical results of this work are considered as a new guide-framework in industry for configuring Cisco MARs and it is recommended to be tested, debugged, HW/SW co-design programmed and developed on updating the commercial products of Cisco, ex. 3200 Series MAR.

REFERENCES

ARTIS Software Corporation. 2001. *Extensions to Petri Nets*. Telecom and Datacom Engineering Solutions, USA.

Bing, B. 2000. *High-Speed Wireless ATM and LANs*. Artech House, Boston, USA, pp. 19, 64-66, 133-148.

Chan, C.W. 2000. *Performance Analysis of Telecommunications and Local Area Networks*. Springer, USA, Equilibrium Point Analysis (EPA): pp. 414-418.

Chodorek, R.R. and A. Glowacz. 2002. "Behavior of TCP Westwood in Wireless Networks". In *Proceedings of ATAMS'2002 International Conference* (Cracow, Dec.). AGH University of Technology, Poland.

Cisco Systems Inc.. 2009. *Cisco 3200 Series Mobile Access Router SW Config Guide*. Revised Jul. 2010.

Dip B. 2006. "Wireless Usage Analysis and Data-Modeling Using a Performance-Based Approach". In *Proceedings of ICNS'2006 IEEE/Cisco International Conference on Networking and Services* (Silicon Valley, Santa Clara, California, USA, Jul). IEEE Computer Society.

EL-DDIN, H. 2007. *New Communication Methods in Wireless Mobile Computer Networks*. PhD. Dissertation, Slovak Technical University, Bratislava, Slovakia (Oct).

Fisher, W. and K. MEIER-HELLSTERN. 1992. "The Markov-Modulated Poisson Process (MMPP) cook-book". In *Performance Evaluation 18*, pp.149-171.

Hodge, J.R. 1996. "A Framework for Simulating Wireless Mobile Client-Server Computing". Master thesis, University of Waterloo, Waterloo, Ontario, Canada.

Jamalipour, A. 2003. *The Wireless Mobile Internet: Architectures, Protocols and Services*. John Wiley & Sons Ltd., England, pp. 247-262, 368-382.

Ozcan, A. et. al. 2010. "Recent Trends in Wireless and Mobile Networks". In *Proceedings of 2nd International Conference WiMo'2010*, Ankara, Turkey (Jun).

Perkins, C. (Ed.) 2002. "IP mobility support for IPv4". In *http://www.rfc-editor.org/rfc/rfc3220.txt*, Revised Oct.2010 (Jan)

Philippart, C.N. 2003. *Markov Modulated Chains*. PhD. dizzertation, Instituto Superior T énico, Lisbona, Portugal.

Raab, S. et al. 2005. *Mobile IP Technology and Applications*, Cisco Systems Inc., Cisco press, USA, Ch. 1, 2 , 4, 7.

AUTHOR BIOGRAPHY

HOSSAM EL-DDIN is from Alexandria, EGYPT. He obtained the B.Sc./Eng.: in „Electronics and Electrical Commun--ications Engineering (EECE)", 1997; Dipl.: in „Information Technology and Computer Networks (IT & CNs)" at ITI-Egyptian CABINET in Cairo, 1998; M.Sc.: in EECE, 1999; defended MPhil.: in „Optical Computer Networks (OCNs)", 2002.. all in EGYPT; and finally hold the PhD. degree: in „New Communication Methods in Wireless Mobile Computer Networks (MCNs)", Governmental Scholarship, at Dept. of Computer Networks and Systems UPSS, Faculty of Informatics and Information Technologies FIIT, Slovak Technical University STU in Bratislava, SLOVAKIA, co. University TU-Vienna, AUSTRIA, 2007.

His C.V. record, so far: 43-IT training courses, 5-scholarships/grants, 7-years residences in EUROPE, 5-years Visa to USA, 4-languges command, 16-publications in IEEE/Cisco/Military proceedings and journals in SLOVAKIA, SINGAPORE, AUSTRIA, TURKEY and USA, 3-citations, 3-European funded projects, 17-conferences, seminars & presentations, and 4-societies' memberships; while he is certified iBK-Vienna AUSTRIA, CEU-IACET USA, and CCNP; and is a Reviewer/Editorial Natural and Applied Sciences committee of World Academy of Science, Engineering and Technology (WASET).

His professions include: „Commercial Teaching Instructor" at ITI-Egyptian CABINET/Cairo, „University Teaching Assistant"-EGYPT 1997-2003; „L2 Technical Support Operations Engineer for EMEA (Europe, Middle-East and Africa)" at IBM/Lenovo Int. in Bratislava SLOVAKIA, „Scientific Researcher" at University TU-Vienna in AUSTRIA, contributions in research projects of Slovak Science Grant Agency with Cisco wireless labs at FIIT-STU and „Assis. Professor" at University STU in Bratislava, SLOVAKIA, engaged with 2-future European funded research projects in CZECH REPUBLIC, 2003-2011; while his research interests include Architectures, Protocols and Algorithms to cope with All-OCNs and Wireless-MCNs.

Dr. Dipl. Eng. Hossam EL-DDIN, *PhD.* has been certified „Achievement Award in Computer Science"-AUSTRIA; Candidate: „Scientist Year STU 2008" and finally awarded „Honorable Recognition for Qualified Research"-SLOVAKIA.

- Staff at STU: http://is.STUba.sk/lide/clovek.pl?id=41639
- Success Story: http://www.iti.gov.eg/SuccessStories.asp

MEDIA ON THE GO

Mobile Advertisement in Vehicular Ad-Hoc Networks

Ciprian Dobre and George Cristian Tudor
Department of Computer Science
University POLITEHNICA of Bucharest
Spl. Independentei, 313, Bucharest
Romania
E-mails: ciprian.dobre@cs.pub.ro, george.tudor@cti.pub.ro

KEYWORDS
Advertisement, data dissemination, wireless environment, modeling and simulation, vehicular ad-hoc networks.

ABSTRACT
Mobile Advertisement is a location-aware dissemination solution built on top of a vehicular ad-hoc network. We envision a network of WiFi access points that dynamically disseminate data to clients running on the car's smart device. The approach can be considered an alternative to the static advertisement billboards and can be useful to business companies wanting to dynamically advertise their products and offers to people driving their car. The clients can subscribe to information based on specific topics. We present design solutions that use access points as emitters for transmitting messages to wireless-enabled devices equipped on vehicles. We also present implementation details for the evaluation of the proposed solution using a simulator designed for VANET application. The results show that the application can be used for transferring a significant amount of data even under difficult conditions, such as when cars are moving at increased speeds, or the congested Wi-Fi network causes significant packet loss.

1. INTRODUCTION

Intelligent Transport Systems (ITS) are formed by adding information and wireless communication capabilities to transport infrastructure and vehicles. Such systems have the potential to increase safety and comfort of drivers (Tarnoff, *et al*, 2009). ITS applications can lead to cooperative collision warning, to congestion-avoidance and the finding of faster and safer routes for drivers. A different category is represented by the non-safety (comfort) applications. Examples include Electronic Toll Collection (ETC) (Lee, *et al*, 2008), car to home communications (Dkesson&Nilsson, 2002), travel and tourism information distribution (O'Grady&O'Hare, 2004), etc.

In this paper we present a solution for the dissemination of information to interested drivers using ITS capabilities. *Mobile advertisement over VANET* is an application that takes advantage of short-range wireless network communication for making recommendations based on location-awareness. It uses Access Points (AP) as emitters, constantly sending messages to all vehicles within range. The main advantage of using short-range wireless technology for this application is that it can exploit the natural locality of the signal. The information is received by vehicles located within the neighboring area. Therefore, by using a short-range signal, the difficult task of finding vehicles within a geographical region is avoided.

Using WiFi communication is difficult because of the cars' speeds. Cars enter and leave the wireless transmission range of an AP at all times, so the window for transmitting the data is low. However, by directly incorporating data into the WiFi beacons we manage to solve this. We show results proving the capability of such an approach to cope with increased car's speeds.

The messages range from commercial advertisements to information about road conditions or traffic congestion. The client application runs on the driver's smart phone, and it receives messages encapsulated in the beacon data sent over the wireless protocols by neighboring access points. Based on the driver's specific preferences, the client is capable to filter and recommend specific information to the end-user.

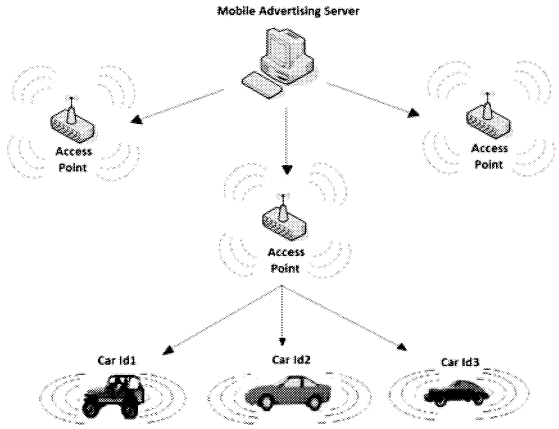
Such an application has the potential to replace the static and more expensive advertisement traditional billboard panels. There are advantages for both the driver, as well as the business provider. The driver is presented with less information, which is more relevant to its specific needs. The business provider advertises its products in a more efficient way. Typical scenarios include a restaurant dynamically advertising its menu or a store advertising its discounts. Or the administrator of a road decides to do some maintenance work on a specific portion. In this case he/she could place APs at locations leading up to the affected area, informing drivers of the hardened driving conditions and possibly offering advice about alternative routes.

In this paper we present the architecture and implementation details of the proposed solution. We present details describing the use of APs for mobile advertising. We outline the obstacles on the implementation and proof of the theoretical expected performance. We also present evaluation results of the proposed solution using modeling and simulation. For that we used a simulator equipped with a realistic mobility model, with streets, routes and various driver behaviors, as well as having a network model including wireless communication protocols adequate for the use in vehicle ad-hoc networks.

The rest of the paper is structured as follows. Section 2 presents the solutions for mobile advertisement in vehicular ad-hoc networks. In Section 3 we present implementation details of a prototype solution. In Section 4 we present experimental results that demonstrate the capabilities of the presented solution. Finally, in Section 5 we give conclusions and present future work.

2. MOBILE ADVERTISING USING ACCESS POINTS

The use of access points for implementing mobile advertisement is appealing because the wireless signal is confined to a specific area. In this case the information is transmitted only to vehicles located in the same geographic region with the AP. In many cases the information is of interest only to vehicles situated in the vicinity of the data source. The information might be, for example, the advertisement of a local sale or a special discount in a restaurant. The solution does not require a discovery phase, and, therefore, it reduces the time needed for the propagation of the information.



Figures 1: System architecture.

In addition, we wanted to develop a solution that can work in a mobile dynamic environment formed between the AP and vehicles constantly changing their positions. According to the 802.11 standard (which regulates the use of wireless local area networks), the nodes must establish a connection prior to any actual communication (Wong, *et al*, 2006). In case of the proposed mobile dissemination solution this is not a viable solution. In this case vehicles are moving at relatively high speeds (50-60 km/h in a typical urban scenario) and, considering that most APs have a limited range of transmission (around 100 m), the cars would not stay in the communication area long enough for a connection to be established, let alone send the required information. Also, a standard AP can only support a limited number of connections at one time, thus causing problems in case of dens traffic. Vehicles would have to send packets back to the access point, thus leading to possible packet collisions and increased packet loss ratio. Any vehicle could only receive messages from one AP at a time.

Because establishing a connection between nodes is not a viable solution, we had to work around the limitations of the 802.11g standard. We propose using beacon frames for carrying the data. According to the wireless protocol, such beacon frames are periodically sent by APs to advertise their presence over the network. These frames can be modified to carry additional information besides the regular network control parameters.

Figure 1 illustrates the general design of the mobile advertising system. Messages to be broadcasted are managed by a mobile advertisement server, which can be connected to several access points. The server passes the messages to the APs, which broadcast them over the wireless network. Vehicles within each AP's wireless transmission range, equipped with wireless devices, receive and filter messages. Based on the user's preferences, a mobile client can further discard or present the message to the driver. For beacon collision avoidance we consider solutions described in (Sgora, *et al*, 2009).

The business provider can dynamically change the advertisement message sent by each individual AP. It can also dynamically manage operational parameters such as the beacon interval, or the advertised wireless network name. Modifying the beacon interval is useful for creating an energy efficient system. As presented in the next Section, data throughput is proportional with the interval between beacon frames. A smaller interval can result in more data being transmitted. For power saving the advertising server can set the beacon interval to a higher or smaller value, depending on the size of the message being transmitted.

The name of the wireless network is used by cars to identify the originator of the received message. In standard wireless networks this parameter should uniquely identify an AP and is set by default to the MAC value. For our solution we took advantage of this field to increase the system's performance. If more APs share the same network name a vehicle passing within range could not differentiate between them. To the vehicle it would appear as if it receives a constant flow of information from a single source. Considering that the wireless range of an access point is limited to about 100 meters, this technique can be useful for increasing the area over which a message is transmitted. A vehicle can seemingly roam from one AP to another without knowing that the source of the message changed by placing multiple APs in such a way that no gap in the wireless signal exists between them.

The solution is based on a push model of information delivery. The emitter keeps sending data without expecting a response. The idea is to overload the 802.11g beacon frames with additional information that can be interpreted and used by the receivers. The beacon frames are sent by APs at fixed intervals to make their presence known over a Wi-Fi network. These frames are received by all clients in range, regardless whether they are connected or not to the AP. A client can also simultaneously receive beacon frames from multiple APs. For that it periodically scans all channels to receive incoming messages encapsulated within beacons.

The custom information is added to a beacon frame by altering its structure and overriding one or more fields. There are *three fields* that can be used (Wong, *et al*, 2006). The **SSID field** carries the name of the wireless network. It has a length of 32 bytes. The advantage of using this field is its relatively easy way of manipulation. Most commercial APs provide a method for setting the SSID field. Clients running different operating systems can get the SSID from a beacon directly in user-space, without requiring specially-adapted drivers for the wireless card. The disadvantage is that beacons carrying useful information would overlap with other beacons sent from access points not participating in our solution. Clients might be presented with information

about fake networks from non-participating APs. The **BSSID field** (6 bytes long) is the unique identifier of the AP. This overcomes the limitation of using the SSID field, but the small size is not practical for our application. Finally, the **Vendor Specific field** can also be used. The 802.11 standard allows AP vendors to add 253 bytes of information at the end of a beacon frame. The use of this field requires updating the AP's operating system and network stack to add vendor-specific information dynamically in the beacon. The wireless network card driver on the client's side must also be updated to allow reading this field and passing the information to user-space applications. However, vendor specific extensions for reading the fields are currently being developed for modern mobile operating systems (Android, iOS) (Berg, 2009).

The alternatives are not enough to transport enough data in the wireless beacon. Therefore, we further considered solutions to send larger chunks of data. We turn to splitting the messages in smaller data chunks that can fit inside a beacon frame. Each frame would then contain a fragment of the message. Each fragment receives an index. The first and last fragments also contain special delimiters for the limits of the message. The AP then keeps sending the message in a loop. If a client happens to lose the beacon message because of the outside interference, it simply notices that the received fragments are not sequential. In this case it can either discard the fragments or, if possible, reconstitute the message from a second loop of the same message.

In our experiments we considered the Vendor Specific field of being capable of carrying the information. It is the one that offers an acceptable bandwidth for fast information transmission. From the 253 bytes available in this field, we use 3 bytes for control and chunk ordering and the other 250 for transmitting the data. The proposed structure of this custom Vendor Specific field is presented in Figure 2.

2B	1B	250B
Seq. No.	First/Last Tag	Data

Figures 2: Vendor Specific fields.

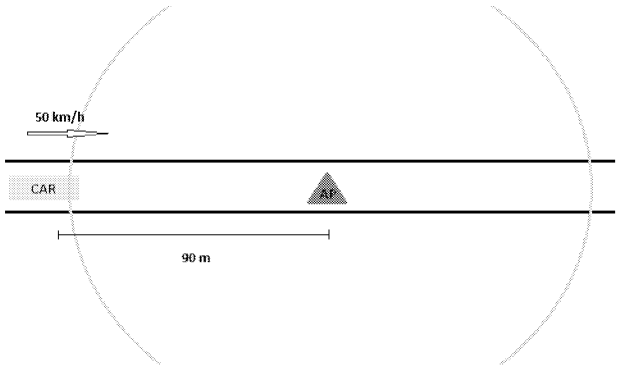
The *Sequence Number* field is two bytes in length and contains the index inside the message for the currently transmitted data fragment. If we were to use only one byte for this field, our message size would have been limited to 256 fragments of 250 bytes each, leaving us with a maximum message size of 62.5 KB.

The *First/Last Tag* field contains *1* for the first fragment of the message and *2* for the last fragment. All other fragments set this field to *0*. Finally, the *Data* field contains the custom data.

An AP periodically sends beacon frames to make its presence known over the wireless network. Under normal operation these frames carry only the communication parameters advertised by the AP. The default interval between beacons for most APs is set to 100 ms. The time interval offers a good balance between power consumption and responsiveness when scanning the medium for discovering the AP. The interval can be manually set to any value between 1 ms and 65535 ms. In our experiments we considered a beacon interval of 10 ms.

The wireless transmission range depends on the 802.11 version and the AP's producer. Several companies claim outdoors ranges of up to 150 m using the 802.11g protocol. However, we used a 90m range, based on previous research results (Singh, *et al*, 2002). Beyond this limit all packets are assumed lost. Nonetheless, the declared maximum range is that at which the AP is designed to be able to maintain a relatively stable connection with a client. However, we do not need establishing a connection, so the real distance at which beacons are received might be greater, even though the loss ratio increases dramatically with the distance.

Based on the assumed figures, we evaluated the maximum amount of information that can be sent. We used a simple model, with an AP on a road, and several cars (see Figure 3).



Figures 3: Car approaching an AP.

Considering a speed limit of 50 km/h for a vehicle travelling in the city, the maximum time it takes before it reaches the access point is equal to $90m/(50km/h) = 6.5s$.

A car would have to receive the message in less than *6.5* seconds, before reaching the AP's position. Considering a beacon interval of *10* ms, the AP is able to send *100* packets every second. In this case the vehicle can receive maximum *650* packets before reaching the AP's location. By adding *250* bytes of data in every packet, the amount of information is $650\,packages \cdot 250\,bytes/package = 158\,KB$. So, the maximum amount of data that can be transmitted before the car reaches the AP is *158* KB, and the total amount of data the car is able to receive before completely exiting the transmission range of the AP is *316* KB.

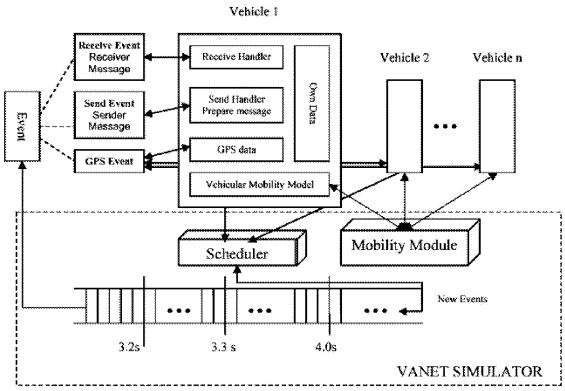
3. IMPLEMENTATION DETAILS

We evaluated the mobile advertisement solution using modeling and simulation. We implemented the solution as an extension on top of a realistic traffic simulator called VNSim (Gainaru, *et al*, 2009).

VNSim is a generic VANET traffic simulator incorporating microscopic and macroscopic traffic, mobility and networking models to accurately evaluate the performance of a wide range of VANET technologies. It is designed as a

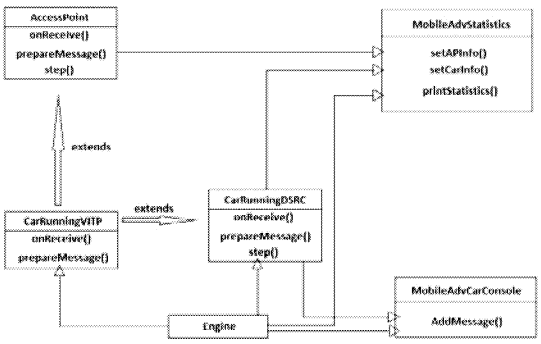
realistic simulator for evaluating the performances of a wide-range of VANET technologies, ranging from wireless networking protocols and dissemination strategies to applications being developed over VANETs.

The simulator is implemented in Java and uses a synthetic mobility model that integrates both microscopic and macroscopic motion. Its mobility model is in charge of importing the map topology and building the dynamic of all vehicles. In addition, the simulator uses a wireless networking model, responsible with the simulation of the networking components and the communication protocols envisioned by a VANET system (see Figure 4).



Figures 4: The architecture of VNSim.

For evaluating the proposed Mobile Advertising solution, we extended VNSim with new components and capabilities. On top of the mobility and network models we added new components to simulate behavior and characteristic of APs and cars. Figure 5 illustrates how the mobile advertisement components are integrated with the VNSim simulator. For example, the engine of the simulator maintains a list of cars. We extended it also hold the APs involved in the simulation experiment. Both cars and APs are extension of the same base class CarRunningVITP (Dikaiakos, *et al*, 2005). The engine creates an instance of the car console and an instance of the statistics class. All the cars and the APs populate these classes directly. At the end of the simulation, the engine displays the statistics.



Figures 5: Implementation details.

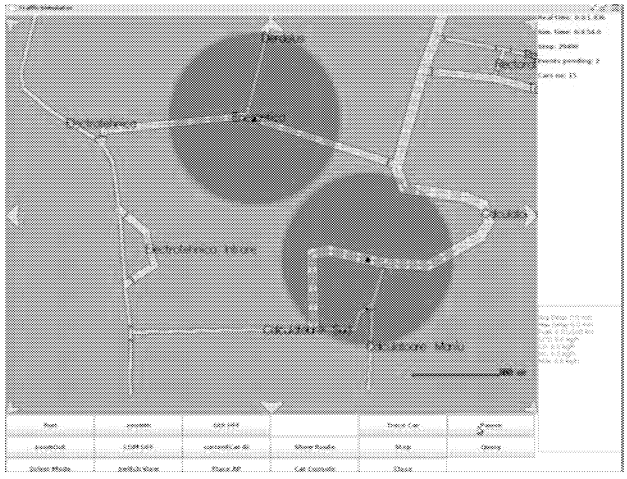
The extended car receives beacon messages and handles them accordingly. In addition, we introduced a new type of message for caring mobile advertisement (the simulation of the proposed beacon message). The simulated car stores the received messages for a certain period of time. Each car maintains a hash map with the keys being the unique identifiers of the APs sending the messages and a structure holding the received information. When a car receives a beacon frame, it extracts the useful data, and, if it is a new fragment, it further stores it.

Another component added in the simulation is an access point that sends message beacons. The position of the AP on the digital map is set by the user using the graphical user interface. The AP is initialized with default values set for beacon interval, message size and range, and it generates a random message to be sent. The user can specify various parameters for the AP, such as wireless transmission range, size of the transmitted messages, etc.

To simulate a noisy environment we also added a function for randomly dropping packets sent by the AP. The user can specify a probability for the loss of packets.

The output of the simulation experiment consists in the average time required to send a message, the number of message loops needed for a car to receive a complete information in a noisy environment, and the ratio of dropped and received messages.

For the output we added a component that logs and displays statistical information. The statistics contain information about the APs (range, beacon intervals, message size, frames sent, complete loops and time running), about individual cars (completed messages, dropped messages, received frames, duplicate frames and lost frames) and general messaging information (total frames sent, total completed messages, frames received per car, frames lost per car etc.).



Figures 6: The extended graphical user interface presenting a running experiment.

We also added additional functionality to the graphical output of VNSim (Figure 6). We extended the interactive simulation mode and included a visual representation of the access points and other components used by the mobile advertisement application.

4. EXPERIMENTAL RESULTS

We present evaluation experiments designed to determine the maximum throughput of the application. The scenario assumes an AP being placed on an open road, without any intersection nearby and without congested traffic. This represents ideal driving conditions in which cars move with relatively constant speed, without having to slow down or stop. The alternative would be to obtain artificially improved results using reduced car speeds, but this would extend the time cars are within wireless range of the AP.

In the experiments the wireless transmission range of the AP was set to 90 m (and remained constant through the simulation experiment). The beacon interval was set to 10 ms. In an experiment we assumed a 5% packet loss probability, and in another one the target was increased to 10%. We also sent messages with sizes varying between 16KB and 112KB.

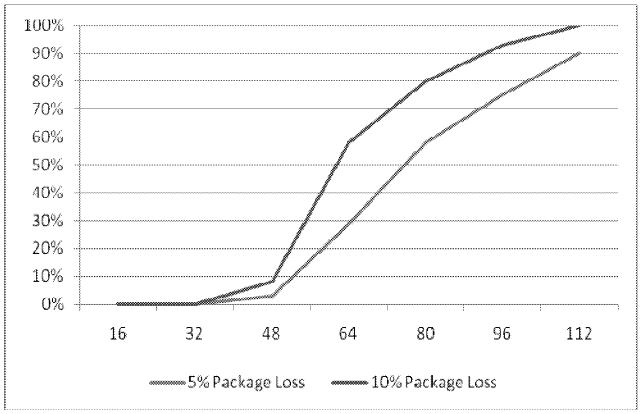
Under these conditions, we executed several experiments using increasing sizes for the messages to be transmitted. For each size we computed a "*Message Loss Percentage*", which indicates the percentage of all cars that did not successfully receive the message. The cars moved with an average speed in the interval 60-70 km/h.

Table 1. Evaluation results.

Packet Loss Probability	Message Size	Message Loss
5%	16 KB	0%
5%	32 KB	0%
5%	48 KB	3%
5%	64 KB	29%
5%	80 KB	58 %
5%	96 KB	75%
5%	112 KB	90%
10 %	16 KB	0%
10 %	32 KB	0%
10 %	48 KB	8 %
10 %	64 KB	58 %
10 %	80 KB	80 %
10 %	96 KB	93 %
10 %	112 KB	100 %

During these experiments we noticed that for messages smaller than 32 KB the message delivery percentage was 100%, even in a transmission medium with only 10% packet loss probability. For messages exceeding 64KB, the ratio becomes unsatisfactory for our application.

Figure 7 presents the results for the *Message Loss*. The ascending slope of the message loss percentage becomes clearer. We notice that, when increasing the medium packet loss probability, not only do performances drop, but they drop at a faster rate.



Figures 7: Evaluation results.

For message sizes of approximately 100 KB the messages are almost completely lost, especially in the case of a medium with 10% packet loss probability. The explanation for such a sudden drop in performance is found when examining the time required for a complete message to be sent and the time that a car running at 60-70 Km/h spends in the access point range.

To exemplify the results, let's consider a message of 112 KB. The number of frames required to send the message (considering that one frame can hold up to 250 Bytes) is 459 frames (112 KB / 250 B). The beacon interval was set to 10 ms. Therefore, sending 459 frames would take 4.59 seconds (459 * 10 ms). The time it takes for a vehicle moving at 70 Km/h to traverse the wireless transmission range of an AP (90 m) is 4.6 seconds (90 m / 70 Km/h). Therefore, it takes a vehicle *4.6 seconds* to reach the AP, and it takes a message *4.59 seconds* to be completely sent. When the medium has a 10% packet loss probability chances are that the message needs more than one loop to be successfully transmitted, but during the second transmission the car leaves the access point range.

5. RELATED WORK

The world-wide expansion of mobile networks placed mobile technology into the hands of millions of people. Innovative programs have been put in place to disseminate crucial health, social and political data over mobile devices and to use them to collect eyewitness reports and personal health information. Mobile phones can communicate across a variety of online platforms and networks, with high-end phones replicating the capabilities of desktop computers. The primary advantage of mobile data dissemination is in using handheld technology to directly reach the intended recipients.

The problem of data dissemination was approached in many cases for the particular case of Vehicular Ad Hoc Networks. Numerous local incidents occur on road networks daily, many of which may lead to congestion and safety hazards. If vehicles can be provided with information about such incidents or traffic conditions in advance, the quality of driving can be improved significantly in terms of time, distance, and safety. An analysis of several solutions is presented in (Sutariya& Pradhan, 2010).

AdTorrent, an integrated system for search, ranking and content delivery in car networks, previously proposed the notion of Digital Billboards, a scalable "push" model architecture for ad content delivery (Nandan, *et al*, 2006). The mobility model for the urban, vehicular scenario can be used in conjunction with the analytical model for estimating query hit ratio by a system designer to determine the scope of the query flooding as a function of the available storage per vehicle for their application.

Zhao *et al* (2007) propose a solution which considers information source (data center) to disseminate data to many vehicles on the roads. It is noted that periodically pouring data on the road is necessary since vehicles receiving the data may move away quickly, and vehicles coming later still need the data. Caliskan *et al* (2006) focuses on the decentralized discovery of parking places. The proposed model consists of communication between vehicles and fixed infrastructures named as parking automat and also between vehicles.

Unlike previous solution, we propose an approach that takes advantage of short-range wireless network communication. As far as we know it is an idea different from previous solutions by at least: the ability to transfer data without establishing a connection beforehand and the fact that the information is confined to the geographical area in which it could be useful.

6. CONCLUSIONS

Mobile advertisement is an application developed on top of a wireless communication infrastructure that provides vehicles on the road with advertisement information related to their current location. The application can be useful to business companies because they can dynamically advertise their products and offers to more people driving their car. We presented a solution for implementing such a system, based on the idea of using access points as emitters for transmitting messages to wireless-enabled devices equipped on vehicles.

Such a solution provides a practical and inexpensive approach for delivering information to vehicles with wireless devices. We presented the approach taken for disseminating information using the beacon frames available in the 802.11 protocol suite. We also presented details about the implementation of the mobile advertisement components on top. We showed how the simulator can be used for evaluating such applications and showed how the mobile advertisement extended its functionality.

The results of the performed experiments reveal that the mobile advertisement solution can transfer a significant amount of data even in difficult conditions, in which cars are moving at increased speeds, or when a congested Wi-Fi network causes significant packet loss. The evaluation results prove that people can more easily disseminate information and can easily discover data of interest using the infrastructure provided by a VANET environment. The performance does not have a large negatively affect on other communications in VANET.

We are currently working on an implementation of the presented solution, using the Vendor Specific field of the 802.11 protocol suite. The solution will be further evaluated in urban traffic scenarios, but we also plan to extend its applications to in-door context-aware dissemination approaches.

ACKNOWLEDGMENTS

The research presented in this paper is supported by national project: "TRANSYS – Models and Techniques for Traffic Optimizing in Urban Environments", Contract No. 4/28.07.2010, Project CNCSIS-PN-II-RU-PD ID: 238. The work has been co-funded by the Sectoral Operational Programme Human Resources Development 2007-2013 of the Romanian Ministry of Labour, Family and Social Protection through the Financial Agreement POSDRU/89/1.5/S/62557.

REFERENCES

Berg, J.M. 2009. 'WiFi Overview'. Last accessed March 02, 2011, from http://linuxwireless.org.

Dikaiakos, M.D., S. Iqbal, T. Nadeem, and L. Iftode. 2005, 'VITP: an information transfer protocol for vehicular computing', *Vehicular Ad Hoc Networks 2005*, pp. 30-39.

Dkesson, K.-P., and A. Nilsson. 2002. 'Designing Leisure Applications for the Mundane Car-Commute', *Personal Ubiquitous Comput.*, 6, 3 (January 2002), pp. 176-187.

Gainaru, A., C. Dobre, and V. Cristea. 2009. 'A Realistic Mobility Model based on Social Networks for the Simulation of VANETs', in *Proc. of the VTC-Spring 2009 Conference*, Barcelona, Spain, pp. 1-5.

Lee, W.-H., S.-S. Tseng, and C.-H. Wang. 2008. 'Design and implementation of electronic toll collection system based on vehicle positioning system techniques', *Comput. Commun,* 31, 12 (July 2008), pp. 2925-2933.

O'Grady, M. J., and G.M.P. O'Hare. 2004. 'Just-in-Time Multimedia Distribution in a Mobile Computing Environment'. *IEEE MultiMedia* 11, 4 (October 2004), pp. 62-74.

Sgora, A., D. D. Vergados, P. Chatzimisios. 2009. 'IEEE 802.11s Wireless Mesh Networks: Challenges and Perspectives'. In *Proceedings of MOBILIGHT'2009*. pp.263-271.

Singh, J. P., N. Bambos, B. Srinivasan, D. Clawin. 2002. 'Wireless LAN Performance Under Varied Stress Conditions in Vehicular Traffic Scenarios'. In *Proceedings of Vehicular Technology Conference* (*VTC 2002 – Fall*). pp. 743-747.

Tarnoff, P.J., D.M. Bullock, and S.E. Young. 2009. 'Continuing Evolution of Travel Time Data Information Collection and Processing'.T*ransportation Research Board Annual Meeting 2009 Paper* #09-2030 TRB 88th Annual Meeting Compendium.

Wong, S.H.Y., H. Yang, H. Songwu Lu, and V. Bharghavan. 2006. 'Robust rate adaptation for 802.11 wireless networks', In *Proceedings of the 12th annual international conference on Mobile computing and networking* (MobiCom '06). ACM, New York, NY, USA, pp. 146-157.

Sutariya, D., S. N. Pradhan. 2010. 'Data Dissemination Techniques in Vehicular Ad Hoc Network'. *Intl. Journal of Computer Applications* (0975 – 8887). Volume 8– No.10, pp. 35-39.

Nandan, A., S. Tewari, S. Das, L. Kleinrock. 2006. 'Modeling epidemic query dissemination in Adtorrent Network'. In *Proc. of IEEE CCNC*, pp. 1173-1177.

Zhao, J., Y. Zhang, G. Cao. 2007. 'Data pouring and buffering on the road: A new data dissemination paradigm for Vehicular Ad

Hoc Networks'. *IEEE Transactions on Vehicular Technology*, vol. 56(6). pp. 3266-3277.

Caliskan, M., D. Graupner, M. Mauve. 2006. 'Decentralized discovery of free parking places'. In *Proc. of the Intl. Conf. on Mobile Computing and Networking*, pp. 30- 39.

BIOGRAPHY

CIPRIAN DOBRE PhD, is lecturer with the Computer Science and Engineering Department of the University POLITEHNICA of Bucharest. The main fields of expertise are Grid Computing, Monitoring and Control of Distributed Systems, Modeling and Simulation, Advanced Networking Architectures, Parallel and Distributed Algorithms. Ciprian Dobre is a member of the RoGRID (Romanian GRID) consortium and is involved in a number of national projects (CNCSIS, GridMOSI, MedioGRID, PEGAF) and international projects (MonALISA, MONARC, VINCI, VNSim, EGEE, SEE-GRID, EU-NCIT). His research activities were awarded with the Innovations in Networking Award for Experimental Applications in 2008 by the Corporation for Education Network Initiatives (CENIC).

SPOT – COLLABORATIVE CREATION OF LOCATION BOUND TIME LAPSE VIDEOS

Britta Weiland
Arnd Steinmetz
Media Department
Hochschule Darmstadt / University of Applied Sciences
Haardtring 100, 64295 Darmstadt, Germany
E-mail: britta.weiland@gmx.de, Arnd.Steinmetz@h-da.de

KEYWORDS
Mobile location based service, Mobile social network, Video of time based development

ABSTRACT

SPOT is a location based service that records visual change and makes it graspable by enabling users to generate time lapse videos in a collaborative way and also view them. Creating content becomes virtually effortless by using smartphones instead of permanently installed equipment to produce single images of the same motif at geotagged Points Of Interest (POIs), named "spots". Images that are taken at the same spot can be perfectly overlaid by users through providing an image, that was previously taken at the scene, as a transparent camera preview overlay. The images are uploaded to a server and converted into a video that can be adjusted based on user settings and streamed to any client application.

PROBLEM DEFINITION

Since smartphones have become widely available to the public, new types of applications and services that leverage the unique opportunities that these devices offer are entering the market. Smartphones are generally equipped with a variety of location- and orientation sensors, and are connected to the internet while being available to the user virtually everywhere and at all times. This allows to show POIs to users in relation to their current position. Also, by utilizing the device camera preview, displaying POIs in augmented reality views becomes possible. So far, several types of POIs can be found and various media such as text, images and/or video can be attached to locations. For example, Flickr (Flickr) and Google Panoramio (Google Panoramio) offer such features. Unlike these available systems, SPOT enables users to produce a thread of images of one location that shows time-based development of one motif instead of many perspectives of it.

RELATED WORK

Location based services utilizing smartphones are common nowadays, but solely rely on textual data, single images and/or prerendered video. No system deals with collaborative production of location based video. The mobile app Pock-it for example allows users to attach multimedia messages to user-specified locations (Coenen and Steinmetz 2008). Self-Guided City Tour is an application in which tourists can inform themselves at specified POIs through HTML contents (Self-Guided City Tour). Unlike SPOT, this project does not rely on user generated content. In contrast to services that aim to determine the locations of POIs through textual data such as annotations (cp. Serdyukov et al. 2009), latitude, longitude and altitude values are used to specify positions in SPOT. Information about camera orientation is not included as the data delivered by GPS receivers and magnet resonance sensors (compasses) is by no means accurate enough to enable an automatic composition of photographs taken over a long period of time by different people.

A system that has similar components as SPOT and also deals with uploading images at specified locations is PhotoCity. Here, the developers are interested in building 3D models out of single images instead of converting them into videos. To show users a location at which an image shall be taken, an image of the 3D model as it already exists is shown with the area of interest marked orange (cp. Tuite et al. 2010). SPOT instead uses a transparent overlay image to give visual hints about the motif.

INTRODUCTION

Visual changes happen everywhere around us every day. Some changes are rather quick and visible within a matter of hours – some are slower and can even take years to show. These changes, no matter how much impact they have, are often not perceived by people. SPOT records change and makes it graspable. Users can go to locations and view how the scenes and objects that are there have looked in the past

and developed until the present. Further, they can not only view the changes but also record them themselves.

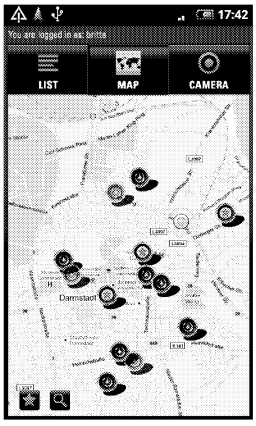
This is done by enabling users to collaboratively create time lapse videos, mainly by using their smartphones. Separate images, which can be taken by different devices and seamlessly overlapped through a transparent overlay of the same motif, are uploaded to a server. These images are then compiled into a video. The resulting films can be viewed in a media player that is also part of SPOT. No permanently installed equipment is necessary to produce content. Therefore, the system makes creating time lapses easier, cheaper and less time consuming than before. Core processes how spots can be edited by users and worked with are the following.

Creating a spot

With the help of a smartphone, users mark a location, take an image at this location and upload it to a database-supported server.

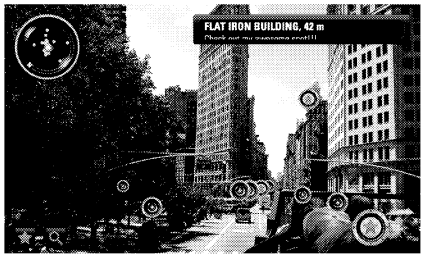
Adding an image to a spot

As shown in Figure 1, users can view where spots are located through (a) a map, (b) list or (c) augmented reality view on their smartphone.



(a) Map

(b) List



(c) Augmented Reality

Figure 1: Location of spots displayed in different views

On either view spots can be selected and the option to add an image chosen, if the spot is within close distance to the user. The image that was last taken for the selected spot is then downloaded from the server and a transparent version of it laid on top of the preview of the smartphone's digital camera. The user then matches overlay image and the image that is shown by the camera preview and presses the shutter release button. After a user confirmation, the taken image is uploaded to the server and added to the spot. When this process is repeated multiple times and optionally by different users, images that show the same motif at different points in time are captured (see Figure 2). Slight shifts in perspective and distance to the object cause a stop motion animation-like character of the video but do not have an impact on perceiving the image sequence as a film that shows a process.

Figure 2: The same object is photographed at different points in time

Viewing a video of a spot

After a spot was selected, the option to view the video of the spot is always available. If chosen, a video that was compiled on the server side out of the images that belong to one spot, is streamed to the user. The appearance and length of this video can optionally be adjusted by the user.

When producing time lapses in a collaborative way, the outcome of the films can not exactly be planned and some may develop in completely unexpected ways. This adds an entirely new aspect to the medium.

By providing users with a way to log in to the system and therefore being able to link data to certain users, SPOT can potentially be extended to provide all assets of a social network. Spots can then be rated, commented on, discussed in forums and more.

ARCHITECTURE

Server

On the technology side, SPOT for one consists of a server that saves metadata about spots and users in a relational database, stores all media files that belong to spots (images as well as video files) and handles client communication.

Utilizing PostgreSQL extended by PostGIS, a spatially enabled database has been created for the system, through which distance calculations between spots can be performed as part of the database query.

Image and video files are saved in a directory structure, in which folders are logically named to enable dynamic creation of paths. To not produce endless amounts of data, a system of strategically saving and deleting data has been developed.

Client communication largely depends on JSON formatted data. Managing file storage, querying the database and sending responses to clients is handled via PHP.

The video conversion process is optimized to allow for various alternative versions of spot videos. To flexibly create these, images of a single spot are filtered and selected depending upon user settings. A H.264 encoded 3GPP video is then created out of relevant images utilizing the tool FFmpeg.

Client

As a first client, a mobile application for the Android platform, optimized for version 2.2 (Froyo), has been developed. As shown in Figure 1 (a), with the help of Google APIs Add-On, Google Maps is embedded into SPOT's mobile client to provide an alternative way of viewing the location of spots. To determine the current location of the user, the device inbuilt GPS receiver is used.

The application enables users to register to the system, log in, create new spots, select spots, add images to them, view the videos in a media player and change the settings at which they would like to view the video.

CONCLUSION

This paper proposes SPOT, a system that provides a novel user experience in the field of location based services for mobile devices. With the help of their smartphones, users can find "spot" POIs, at which images of the same motif can be uploaded to a server and streamed as a video. Initial prototype testing resulted in several threads of images which could successfully be converted into and viewed as videos.

Even though several persons have been introduced to the system and were able to intuitively operate it, no distinct usability testing has been undertaken at this point. Testing the system functionality and usability with a larger group of people and processing the results of these tests is one of the next steps in the development of SPOT.

Generally, the compiled videos have a jittery, stop motion animation-like character and the motif is always slightly moving because of small differences in camera orientation and distance. However, all persons who were presented the videos and told how these were created described the media as impressive.

In summary, it can be said that SPOT is a prototype that picks up the approach to create a unique form of data out of user generated content in a novel way. This prototypical approach is still at its beginning but first tests have been promising and the system has the potential to be expanded into a full-fledged platform in the future.

REFERENCES

Coenen C., Steinmetz A. 2008. "Pock-it: Annotate the world" *Proceedings of the ACM Mindtrek 2008*, 105-108.

Flickr *http://www.flickr.com/map/* Requested on 05. Feb 2011.

Google Panoramio *http://www.panoramio.com/* Requested on 05. Feb 2011.

Self-Guided City Tour *http://www.uni-goettingen.de/de/63850.html* Requested on 28. Feb 2011.

Serdyukov P., Murdock V. and van Zwol R. 2009. "Placing Flickr Photos on a Map" *Proceedings of the ACM SIGIR 2009*, 484-491.

Tuite K., Snavely N., Hsiao D.-Y., Smith A.M. and Popovic Z. 2010. "Reconstructing the World in 3D: Bringing Games with a Purpose Outdoors." *Proceedings of the ACM SIGKDD 2010 Workshop on Human Computation*, 41-44.

BIOGRAPHY

BRITTA WEILAND has studied at Western Michigan University, MI, USA, Cork Institute of Technology in Cork, Ireland and the University of Applied Sciences in Darmstadt, Germany and obtained her degrees in Media Production (B.A.) and Media Direction (M.A.) in 2008 and 2011. She has been active as a freelancer in the fields of Web Development and User Interface Design since 2007.

ARND STEINMETZ is professor at the University of Applied Sciences in Darmstadt, Germany. He is head of the faculty of media and director of the Institute of advanced Learning, Media and Simulation. His research interests include mobile and location based systems, long term archival of digital documents, physical computer interaction. He is member of SIGMM, and German GI.FG MCI.

MEDIA INTERPRETATION SYSTEMS

Biometric Passport Border Security and Holder Privacy Issues and Recommendations

Hasan Alsaedy
Gilles Richard
British Institute of Technology and E-commerce (BITE)
Avicenna House, 258-262 Romford Road E7 9HZ London, UK
E-mail: {hasan,grichard}@bite.ac.uk

KEYWORDS
Biometrics, passports, security, privacy.

ABSTRACT

Biometric documents as passports and identity cards are now available for use to US, European and non European citizen. The core technology in use in these documents is the Radio Frequency IDentification (RFID) technology. Data stored in RFID tag can be remotely read and communication channel between tag reader and the document tag could be intercepted. Reading and intercepting the biometric data is a clear threat to nation border security and document holder privacy. In this paper we investigate the current technologies, we point out their drawbacks and we propose more satisfactory alternative solutions.

INTRODUCTION

It is obvious that the 11th of September is the reason behind the introduction of the US and the UK biometric passport and Identity card with as main objective the border security of the US and European nations. This technology is nowadays largely implemented or in the verge to be implemented in a vast majority of European countries. Nevertheless, human right activists are generally against the introduction of this biometric item. Their main argument is that this technology will violate the privacy of the holder, privacy which is protected by the constitutive status of the states. They believe that not only international gangs but also governments, secret police are likely to violate the privacy of the biometric passport holder. These passports are based on the recommendations and guidelines of the International Civil Aviation Organization (ICAO), the responsible body for setting the international passport standards. In addition to that, the European Union developed its own standard known as Extended Access Control (Hoepman et al. 2006). US government sets up on its web site a frequently asked question (USFAQ) about the US biometric passport: on this web site, the biometric passport concepts are defined, some features of the passport are displayed and the logo of the passport is displayed. Compromised security, violation of holder's privacy and unjustified high cost are frequent critics of the biometric passport in media. Compromising classified from cloning to skimming and eavesdropping. Cloning is the act of coping data from a proximity tag to another proximity tag. Skimming is the act of reading the data stored in proximity tag without the consent of the legitimate passport holder's. Eavesdropping is the interception of transmitted data between the proximity tag and the tag reader. In this paper the concepts, technology and techniques used in the biometric documents are clarified. Border security against privacy issues are analysed. This paper is structured as follows: in the following section, we briefly recall what the biometric documents are and how they are structured. In section 3 and 4, we provide the technical background allowing to understand the issues related to that kind of documents and how security is provided. In section 5, we describe the technique in use to access the data stored on the document. This raises the main issue of privacy: in section 6, we finally recommend alternative techniques allowing to protect user privacy but still enhancing the level of security.

BIOMETRIC DOCUMENTS

1n this paper, the term biometric document will refer to the biometric passport or biometric identity card or both. A biometric document is the same as a traditional document but with a specific feature; the new added feature is a small integrated data storage circuit, embedded in the document, with Radio Frequency Propagation capability. An RF blocking cage is also used to prevent unauthorized reading of the data in the documents' chip, given that the document is closed. The electronically stored data includes the printed document's data, a digital photograph and a unique identification number. The stored digital data are digitally signed to prevent the modification of the stored data. The stored data could include a lot of biometric data such as the fingerprint and the iris scans of the document holder etc. A dedicated reader is used in airports and security points to read and compare the digital stored data with the printed document's data. On top of that, the digital data could be read remotely by activating the chip using Radio Frequency. This will raise document holder's privacy issue; however using strong encryption algorithm to encrypt the data will mitigate the privacy issue. To make the reader aware of what the biometric document is and what concern it might raise, the associated technologies and techniques need to be explained: this is the object of the next section.

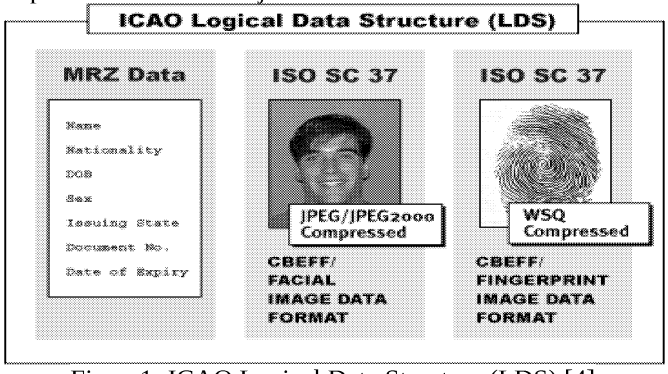


Figure1: ICAO Logical Data Structure (LDS) [4]

TECHNOLOGICAL BACKGROUND

In this section we discuss the technological background underlying biometric documents namely the RFID technology, hashing process and encryption algorithms.

RFID technology

RFID technology is implemented via a tag attached to the target item (product, animal or even person) and a reader allowing to get the information stored on the tag for the purpose of identification and tracking. Depending of the tag technology, it is possible to remotely track the object, the distance depending on the strength of the radio signal. In that case, the tag device contains two parts: the first one is dedicated to the digital storage of the data and the second one is the radio communication antenna. Three types of RFID chips are commercially available nowadays:

- these are the active RFID chips, which contain battery and can autonomously transmit data,
- passive RFID chips, which have no battery and need a Radio Frequency external source to activate signal transmission and
- the battery assisted passive chips requiring an external source of power supply to get activated, the receiving/transmission power of this is significantly high.

These devices are the backbone of what is known as the "Internet of things" where any kind of item can be tracked. For the purpose of biometric identification, it is quite evident that the RFID chip used is the passive one. It is claimed that the RFID chip used in passport is readable only from distance as short as 50 centimeters, whoever some references indicate that the data stream between the RFID device and the RF reader could be intercepted from ten meters distance. The later raises the issue of privacy of the document's holder. Shielding document's cover will significantly reduce the risk. Again encrypting the data transmission will help reducing the risk of the violation of privacy. On the assumption that the document jacket is shielded, the security of the holder will be saved. Opening the jacket of the document will increase the privacy violation risks. In the case of passport, (ISO/IEC 14443) suggests contactless smart card or tag. Distance between the card and the card reader is set as 10cm. The card is powered by the Radio Frequency field of the reader and no battery is needed.

Encryption algorithms

Securing data privacy on electronic storage device and transmission channel needs the usage of encryption techniques and cryptographic algorithms. The stored/transmitted digital data are encrypted by using Public Key Infrastructure (PKI) and Data Encryption Standard (DES) algorithm. Again, the data or the digest of the data could be stored in a dedicated secured database system. The serial numbers of the reported lost or stolen passports are also saved in this database. Symmetric encryption algorithms using only one encryption/decryption key are not as safe as asymmetric algorithms where 2 keys are needed: one key to encrypt and the other one to decrypt. One key is known as the public key and the second one as the private key. PKI is based on these two keys encryption algorithms; each user having a pair of public/private keys. RSA is an example of the two keys algorithm. Web sites of the biometric passport of European countries, non European courtiers and ICAO indicate that RSA and DES are recommended to secure the data stored/transmitted of the biometric passport: RSA is used at the beginning of the communication to exchange the key which will be used to encrypt the data with DES.

Digital Signature

A hash function is a function that takes as input a variable size input binary file and generates as output a fixed size binary string, the generated binary string is the hash value [7]. A good cryptographic hash function should satisfy diverse requirements like no limit in the size of the input file, the output string should be of fixed size, etc.
A well-known example of a cryptographic hash function is the message digest MD5 and the more secure SHA-1[8]. The main use of cryptographic hash function in communications system is to digitally sign data blocks. In this article the hash function is proposed as a tool for digital signature and for calculation of the biometric passport encryption keys.

To digitally sign a document, the hash (or the digest) of the document is calculated and signed by using public key algorithm. To check that the data has not been modified, a other copy of the data is used to calculate the digest and the digest of the signed digest is released. If the two digest are the same, the stored data is accepted, and rejected otherwise.

MACHINE READABLE ZONE COMPLEXITY

The security of the biometric passport is based on the use of the machine readable zone (MRZ): machine readable data are used to generate keys to encrypt and authenticate the biometric data. MRZ consists of four components, these are the passport number (9 digits), the date of birth of the holder (6 digits), the passport expiration date (6 digits) and check digits (three digits one digit for each of the mention three strings). Calculating the complexity of the brute force attack required the estimation of the depth of the search binary tree of the first three components of the MRZ.
The time complexity of the passport number is $9*\log_2(26 + 10) = 46.53$ bits and the time complexity of the date of birth is $\log_2(100*365.25) = 15.16$ bits and the time complexity of the expiring date is $\log_2(5 * 365.25 * 5/7) = 10.35$. The overall time complexity of the brute force attack is 71.86 which is in theory relatively high. The passport holder usually gives willingly copy of his passport to hotel receptionist and foreign exchange shops and many others: accordingly, the MRZ data is available to those people.

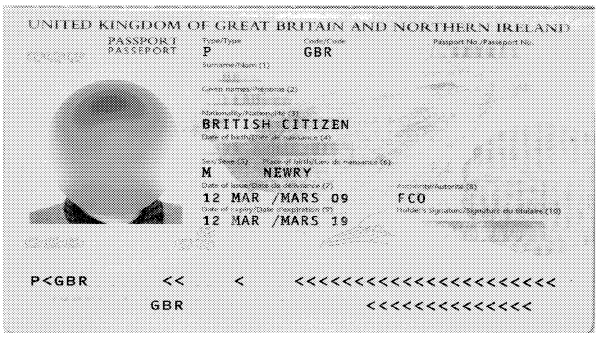


Figure 2: First Page of the UK Passport

Basic Access Control Protocol

In this process, the biometric data stored on the passport are compared with and authenticated with regard to the data stored in the ISO/IEC 14443 tag. The tag usually called

proximity tag. The following steps take place during the document validation. It is assumed that the keys K_{ENC} and K_{MAK} are stored in tag memory during the passport issuing. The following zero knowledge protocol is used.
- The reader read the MRZ from the first page of the passport and generates two keys K_{ENC} and K_{MAK}.
- Tag provides a 64bits challenge sent to the reader.
- Reader encrypts the 64bits challenge using the keys K_{ENC} and K_{MAK} and sends the result back to the tag.
- Tag verifies the response which has to be identical to the encryption of the 64bits using the stored keys. In case of satisfaction, the tag releases its biometric data to the reader.
For the reader to gain access to passport tag, it must demonstrate the knowledge of the access keys (K_{ENC} and K_{MAK}). Calculation of the keys K_{ENC} and K_{MAK} (which are are stored in the tag memory) is based on the following steps:
- *Step 1*: a key seed K_{seed} is derived by taking the most significant 16 bytes of the output of the SHA-1 hash of the MRZ data.
- *Step 2*: to calculate the encryption keys KENC, the K_{seed} are concatenated with 1 and hashed using SHA-1 to 20 bytes digest. Byte 1..8 form the first DES key and bytes 9..16 form the second DES key. Parity bits are adjusted to form the final two 3DES keys. 64 bit 0 string is used as initial value to calculate the value of K_{ENC} using 3DES in CBC mode.
- *Step 3*: To calculate K_{MAK} K_{seed} is concatenated with 2 and the same calculation specified in step 2 is used.

RECOMMENDATIONS AND CONCLUSION

Given that the used technology is commercially available and could be imported from manufacturers world wide, it becomes clear that the national and international border security is not enhanced as much as it should be. And, on the other hand, the privacy of the passport holder is violated unjustifiably. Below is a list of identified issues:

- Passport with defected tag will be problem because access will be denied everywhere to the document's holder. UK Border Agency suggested that passports with defected tag should be processed as a traditional passport. This raises the problem of deliberately damaged tag.
- The privacy of the passport holder is violated as the propagated data from the tag could be traced by secret policy, international gangs and others for many reasons
- Passport data could be read/intercepted from distance during the passport reading process by authorized reader.
- Storing the encryption key on the MRZ data is problematic since MRZ is freely available to external people as the hotel receptionist.
- Finally, the cost of the new passport is higher than the cost of traditional passport.

No contactless tag/USB-like connector
As previously suggested in (Hoepman et al. 2006), we strongly think that the contactless tag must be replaced by a contact tag *with no wireless communication possibility.* There is as well the possibility to communicate via a usb connection by including a small usb connector to the passport (this is under investigation with (ISO 7816-12)). This will eliminate the public concern about the possibility to remotely read the data.

Remember that this is not only a matter of privacy: it can also be a matter of security. If some malicious third party is able to remotely read your private data and then to get your nationality, they can make a guess about the wealth of your relatives. For instance, if you are a citizen coming from a rich country, the likelihood to get money from your relatives to release you from a kidnapping is very high...This could give some ideas to some gang. So it is desirable to protect your citizenship from remote readers.

Database connection
Its has been suggested in (Hoepman et al. 2006) to implement an online terminal authentication to assess the access rights of the reader. Despite this is a good point, we think that the fact that the encryption keys are generated from publicly readable data allows any third party aware of the technology to generate the keys, then to successfully answer the challenge and get the private data delivered by the tag.
We suggest another way to proceed: the reader will have to communicate with a central database server containing a hash of the private data stored on the tag. When the tag delivers the data to the reader, the reader sends a hash of the data to the server which will compare the stored hash with the incoming hash. In case of accordance, the passport is considered as a valid one. In some sense, we use a third party server to authenticate the document. This server has to be maintained by the border security agencies.

More sophisticated hash function
As it is well-known, a hash function comes with a certain percentage of collision which could compromise the system and therefore violate the boarder security and the holder privacy. With the increasing number of travelers, the probability of collision seriously increase. Hence, why not use perfect ones (i.e. collision-free) whatever their cost? It will certainly make the system safer.

We think our method could drastically improve the security of the biometric documents till securing the document's holder privacy. Other options are available but there is a balance to find between the safety of the process and its complexity, to avoid long queue of travelers at the borders!

REFERENCES

J.-H. Hoepman, E. Hubbers, B. Jacobs, M. Oostdijk, and R. Wichers Schreur. 2006. “Crossing Borders: Security and Privacy Issues of the European E-passport,” Proc. First Int’l Workshop Security (IWSEC ’06), 152-167.

Ari Juels, David Molnar and David Wagner, 2005 “Security and Privacy in E-Passport”, http://eprint.iacr.org/2005/095.pdf.

ICAO, 2004 “Machine readable travel documents”. Doc. 9303.

ISO 14443, An introduction to the contactless standard for smart cards and its relevance to customers, http://www.otiglobal.com/objects/ISO%2014443%20WP%204.11.pdf

ISO 7816-12. ISO/IEC 7816 “Identification cards – Integrated circuit(s) cards –Part 12: Cards with contacts – USB electrical interface and operating procedures”. Technical report, ISO JTC 1/SC 17

WEB REFERENCES

The U.S. Electronic Passport Frequently Asked Questions, http://travel.state.gov/passport/eppt/eppt_2788.html

John Leyden, 2006. “e-passport cloning risk exposed, RFID hack attack” http://www.theregister.co.uk/2006/08/04/e-passport_hack_attack , 14/8/2006.

Automatic face analysis for affect recognition

Dragoş Datcu
SeWaCo
Faculty of Military Sciences
Netherlands Defence Academy
The Netherlands
D.Datcu@nlda.nl

Mirela Popa
Philips Research
Eindhoven, the Netherlands/
Man-Machine Interaction group
Delft University of Technology
M.C.Popa@tudelft.nl

Léon J.M. Rothkrantz
SeWaCo
Faculty of Military Sciences
Netherlands Defence Academy
The Netherlands
L.J.M.Rothkrantz@tudelft.nl

KEYWORDS
Face analysis, 3D Active Appearance Model, Facial expressions.

ABSTRACT

Emotions play an important role in every-day life of human beings. Researching the way automatic face analysis can be done represents an important factor in the development of various user-friendly systems and interfaces. In this paper we present the results of our research in automatic recognition of affect state of the users based on video analysis of faces. The approach makes use of Viola&Jones face detection, Active Appearance Models for face shape and texture extraction and Adaboost.M2 for recognition of affects. The novelty of the work consists of a new algorithm for 3D video face analysis. AAM is used to extract visual features such as appearance-related parameters. These are further on used both ways first to detect facial expressions and second to compute the gaze. The facial expressions and gaze provide indications for the psychological profile of the user given the level of responsiveness and perceptiveness in specific situations.

INTRODUCTION

It is the human nature to estimate a person's psychological state following the observation on his/her face. Non-verbal communication channels are typically set during common interpersonal relations and visual messages are processed in a transparent manner.
Nowadays, computer systems already influence significant aspects of our lives, from audio-visual network communication to means of transportation, industry and production. Based on the scientific knowledge about life and environment, we have constantly attempted to improve the work conditions firstly by reducing the workload of the human factor with the functionality of operational automatic systems. The every-day life benefits now from the availability of systems to ease the access to information and facilitates the assistance of individuals in carrying out different activities.
Currently, the general tendency is to construct automatic systems that are able to understand the environmental world and to interact with the users. Human-computer interfaces play an essential role for the generation of system feedback. In this context, the advantage of making machines to read human facial expressions is tremendous. Facial expressions genuinely reveal emotion characteristics of the expresser. Human computer interaction represents a specific area which employs advancements from the research on automatic analysis of affect and behavior. These adaptations of novel technologies are supposed to help increasing the level of interation between the users and computers. Figure 1 illustrates an example of face detection and computation of face shape information with the purpose of facial expression recognition.
This paper proposes a novel technique for automatic recognition of users' affect by using 3D Active Appearance Models as an essential face analysis component between the step which involves the detection of faces and the step which regards the classification of emotions. The next section presents related works on the 2D and 3D face analysis. The following section describes the theoretical base for the 3D AAM. Then we discuss the procedure for preparing the 2D and 3D databases. The next section illustrates the results of recognizing facial expressions. The last section presents the conclusions and future works of the project.

Figure 1: Face detection

RELATED WORK

The analysis of faces implies the segmentation of image regions that are parts of the face. This is done by using various visual features such as rectangular or elliptical parameters. Alternatively, the face location may be determined using the location of the eyes detected in video data from infra-red camera or by using pixel intensity methods that first search for skin-regions in the image. In several papers, the detection of faces in images has been approached by deformable models which represent the variations in either shape or texture of the face object. The

active shape models - ASM (Cootes et al. 1995) and active appearance models - AAM (Edwards et al. 1998) are two deformable models that have been extensively researched and used with good results in the literature. Point distribution models - PDMs relate to a class of methods used to represent flexible objects through sets of feature points that indicate deformable shapes. Marcel et al. (Marcel et al. 2006) use ASMs and local binary patterns - LBPs to localize the faces within image samples. Tong et al. (Tong et al. 2007) propose a two-level hierarchical face shape model to simultaneously characterize the global shape of a human face and the local structural details of each facial component. The shape variations of facial components are handled using multi-state local shape models. Datcu and Rothkrantz (Datcu and Rothkrantz 2008) proposed a bimodal multimodal semantic data fusion model that determines the most probable emotion. The approach uses 2D active appearance models and support vector machine classifiers.

Edwards et al. (Edwards et al. 1998) introduced the AAM as a method to analyze the objects using both shape and grey-level appearances. The spatial relationships are determined using principal components analysis - PCA that build statistical models of shape variation. In a similar way, statistical models of grey-level appearance are derived by applying PCA on shape-free samples obtained by wrapping the face images using triangulation. The models of shape appearance and grey-level appearance are finally combined using PCA, to derive appearance vectors that control both grey-level and shape data.

3D face models have been proposed as an alternative to 2D face analysis for better handling the face pose and face gestures. Dornaika and Ahlberg (Dornaika and Ahlberg 2006) propose two appearance-based methods that use locally exhaustive and directed search for both simultaneous and decoupled computation of 3D head pose and facial expressions.

Lefèvre and Odobez (Lefèvre and Odobez 2010) present an approach that makes use of view-based templates learned online and an extension of a deformable 3D face model to collect appearance information from head sides and from the face. The method has a high robustness in tracking natural and fake facial actions as well as specific head movements.

MODEL

Active appearance model - AAM is a statistical method that handles shape and texture variations of photo-realistic appearance. Seen as a top-down approach, the AAM makes use of prior knowledge on the grey-level appearance, shape structures as well as their relationships, in order to build generative models for the global analysis of a specific class of objects. In our approach we use AAM for extracting the appearance of user's face. Considering that AAM is based on supervised learning, we make use of data sets of face samples for training and testing the model. A natural start is to employ 3D face samples. Additionally, we increase the variation and the size of the database by including also samples from existent 2D facial expression databases. This is accomplished by mapping 2D face samples from the 2D data set to 3D face shapes from the 3D data set.

Active Appearance Models

Building a face appearance model implies a prior acquisition of the face data set. Each face sample has to be consistently annotated with a set of landmark points. The selection of these points is primarily based on the location of facial feature boundaries, at 'T' junctions between boundaries or other relevant face areas. Figure 2 shows an example of a 3D face sample obtained following the key point selection method presented in the previous subsection. The samples depicted in figure 3 and figure 4 correspond to reduced face models of 111 3D points and 185 triangles. Each sample of face shape is described using n 3D key points (x_i, y_i, z_i) by a shape vector: $x = (x_1, y_1, z_1, x_2, y_2, z_2, \ldots, x_n, y_n, z_n)$.

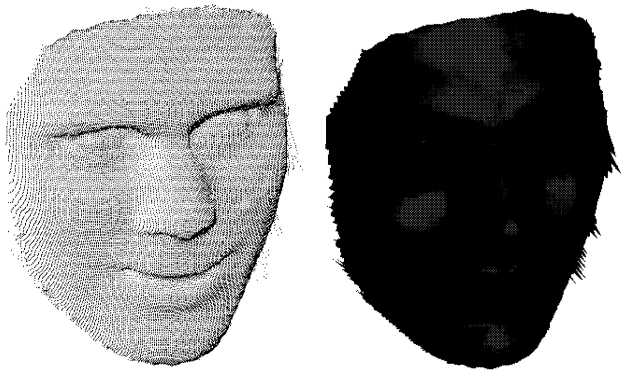


Figure 2: 3D face sample

The face samples obtained were used in the context of applying classic Active Appearance Model on the 3D face data.

The extraction of shape and texture of the face from an image is equivalent to an optimization problem that involves the criterion of minimizing the difference between the real face image and the one generated by the appearance model.

The distance measure can be written as follows:

$$r(p) = g_{im} - g_m$$

where g_{im} is the vector of grey level values of the face patch in the input image and g_m is the vector of grey level values for the face image as it is estimated by the current model parameters p. Alternatevly, vectors g_{im} and g_m can store the color information of the face texture.

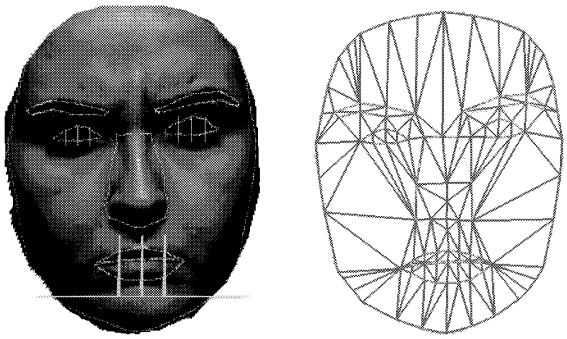


Figure 3: Annotation tool for 3D faces, annotation rules of facial key points

The matching implies finding the optimal appearance model parameters which would lead to the minimization of a scalar measure on the image difference, such as the sum of squares of elements: $E(p) = r^T r$. By considering the computation of eigenvectors on the shape and texture data, the following reconstruction formulae hold:

$$x = \bar{x} + P_s b_s,$$

$$g = \bar{g} + P_g b_g,$$

where $\bar{x}$ and $\bar{g}$ are the mean shape and mean texture, P_s and P_g represent the variation of the orthogonal modes related to shape and texture respectively b_s and b_g represent the shape and texture parameters.

The shape and texture models can be combined into an appearance model as follows:

$$b = Qc,$$

where c is the appearance vector, Q is the matrix of eigenvectors and b is a vector that combines the shape and texture data, as follows:

$$b = \begin{pmatrix} W_s b_s \\ b_g \end{pmatrix} = \begin{pmatrix} W_s P_s^T (x - \bar{x}) \\ P_g^T (g - \bar{g}) \end{pmatrix}$$

, where W_s represents a diagonal matrix to compensate for the different types of data.
The appearance of the 3D face sample vectors can be computed using the formula: $c = Q^T b$.
Given an appearance vector c, the shape and texture of the face can be generated by using the following reconstruction formulae:

$$x = \bar{x} + P_s W_s^{-1} Q_s c$$

$$g = \bar{g} + P_g Q_g c$$

, where $Q = \begin{pmatrix} Q_s \\ Q_g \end{pmatrix}$.

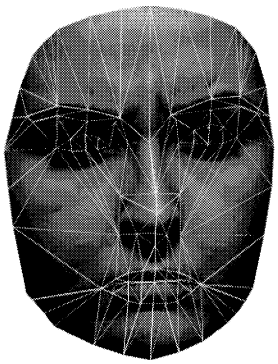


Figure 4: Reduced 3D face sample

The analysis of a new face involves running a fitting procedure during which the appearance parameters are found in such as was so as to minimize the term $r^T r$. This is done by itteratively adjusting the appearance parameters with a term δp:

$$r(p + \delta p) = r(p) + \frac{\partial r}{\partial p} \delta p .$$

The term $\frac{\partial r}{\partial p}$ represents a matrix made up by ij the elements $\frac{dr_i}{dp_j}$:

$$\frac{\partial r}{\partial p} = \begin{bmatrix} \frac{dr_1}{dp_1} & \cdots & \frac{dr_1}{dp_Q} \\ \vdots & & \vdots \\ \frac{dr_M}{dp_1} & \cdots & \frac{dr_M}{dp_Q} \end{bmatrix}$$

The fitting procedure assumes that for a specific current residual r, a p is chosen so that the value of $|r(p + \delta p)|^2$ is minimum. The choice has the form $\delta p = -Rr(p)$, where:

$$R = \left(\frac{\partial r}{\partial p}^T \frac{\partial r}{\partial p} \right)^{-1} \frac{\partial r}{\partial p}^T$$

The term $\frac{\partial r}{\partial p}$ is computed once, during the training step, by altering the elements of p with some amount (typically up to 0.5 standard deviation) and by computing the residuals for the face images included in the training data set.

2D database

Prior to developing models for facial expression recognition, the video database of faces had to be prepared. For training and testing the models, we have used data samples from the Cohn-Kanade database (Kanade et al. 2000). The database contains image sequences from neutral to target display were digitized into 640 by 480 or 490 pixel arrays with 8-bit

precision for grey-scale values. Subjects show various facial displays during different recording sessions. In total, the database contains 486 samples.

3D database

To build the models for automatic 3D analysis of faces, we used the Bosphorus 3D face database (Savran et al. 2008). This database contains 3D face samples showing different facial expressions, variation of poses and certain types of occlusions. All 3D face samples are captured using a structured-light based 3D digitizer device. In order to make use of face samples from this database, we developed specific tools to convert and process the 3D object data.
The original face samples are organized as pairs of 3D point sets and face texture images. Figure 2 illustrates an example of a 3D face sample consisting of 31.032 3D points and 61.359 triangles. Prior to be used for AAM, the face samples had to be converted to a common format which implies the same size of the shape and texture vectors. For that we have developed a tool for manually annotating faces by setting specific key points on the 3D face.
Figure 5 illustrates the 3D cursor of the annotation tool. The cursor can be used to navigate along the 3D geometry of the face and to set face key points. The initial set of key points contains 12 landmarks.

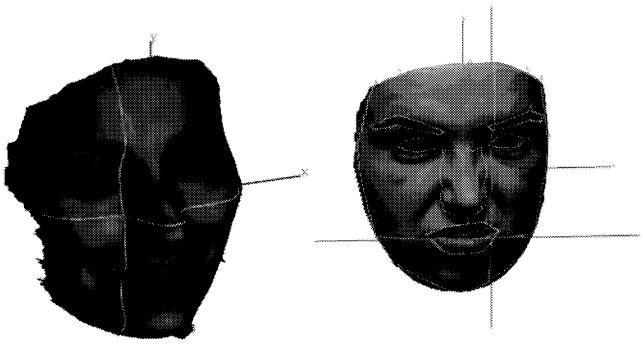


Figure 5: Annotation tool for 3D faces, 3D cursor (left) and face contours (right).

Based on the basic set, the tool automatically computes the points on the 3D segments between the manually annotated key points.
The setting of key points follows basic selection rules according to which some points have to be chosen at given distances, i.e. at the middle the distance between the inner and outer corners of the mouth, on the same vertical line with the inner corner of the eyes, etc.
In order to ease the manual annotation procedure, the tool can display vertical and horizontal guiding planes at specific locations. Further on, it can generate more key points, for instance by considering points at equal distances on these segments or by applying triangulation on the set of key points and by selecting the center of each triangle. The procedure of selecting more points starting from the basic set of landmarks can be stored as a script and can be later used for finding extra landmarks on any other face sample. This sequence of steps ensures for the consistency of points and facets among different face samples in the final data set.

RESULTS

The AAM training data set consists in 317 manually annotated frontal face images. Each face image sample contains 122 landmark points which are located along the shape contours of the face, the nose, the mouth, the eyes and the eyebrows. The face vectors are represented in the shape-free texture space using 45.117 pixel grey-level values.
For testing the AAM models we used a set of previously annotated frontal face images. The testing consists of displacing characteristic shape parameters and in measuring the capability of the model to estimate these changes. All the displacements done for the experiments take as reference point the pixel located at the middle of the line segment delimited by the locations of the inner corners of the eyes.
The model can handle translation on both sides from the true face shape position, up to 20% (about 45 pixels) along X axis and up to 12% (about 27 pixels) along Y axis. In the same way, changes in the face shape scale are correctable in the interval from 0.7 the original face width (about 160 pixels) up to 1.4 the original face width (about 320 pixels). For testing, each face shape has been systematically rotated with angles in the range [-90°,+90°] around the centre point of the line segment delimited by the eye positions. According to the test, the model can adjust face shape rotations of 30° at both sides from the vertical line, around the eyes' centre point.

Recognition of emotions

To address the problem of facial expression recognition from single images, in our approach we extract parametric information with high discrimination power from facial feature space and use it in a data-driven classification environment. We show how to identify such relevant features with the use of boosting methods.
Viola&Jones features have been initially proposed as a fast method for performing object detection (Viola and Jones 2001) and later for face detection (Viola and Jones 2004). For recognizing facial expressions, we used three types of simple features derived from pixel intensities of video frame images.
Local binary patterns represent micro-pattern structural features which can successfully facilitate a reliable representation of the face images. The original LBP operator was introduced by Ojala et al. (Ojala et al. 1996) in the broad context of texture analysis. According to this, a binary pattern at a texture image pixel is generated by applying a threshold on the intensities or grey values of the surrounding pixels in a 3×3 neighborhood with the intensity of the pixel itself. Later, the work of (Ojala et al. 2002) proposed the use of scaled LBP features which allow for sampling from neighborhoods of different sizes. For analyzing the face, we divided the frame image that corresponds to the face into small regions. LBP codes were computed from each image region and concatenated into an enhanced feature vector.

The splitting of the face image is done on 1,2,3,...,10 blocks along X and Y axes with overlap of 0% to 50% with the increment of 10%. From each block we extract features of 4, 6, 8 and 10 neighboring points. In total, the LBP feature extraction produces a set of 72.600 codes for each face sample.

Table 1: The confusion matrix of Adaboost.M2 classifier trained on LBP features and leave-one-out cross validation

	Fear	Surprise	Sadness	Anger	Disgust	Happy
Fear	**80.00**	0.00	0.00	0.00	0.00	20.00
Surprise	0.00	**100.00**	0.00	0.00	0.00	0.00
Sadness	10.00	10.00	**60.00**	10.00	10.00	0.00
Anger	0.00	0.00	20.00	**40.00**	20.00	20.00
Disgust	0.00	0.00	8.33	8.33	**75.00**	8.33
Happy	11.11	0.00	0.00	0.00	0.00	**88.88**

Table 2: The confusion matrix of Adaboost.M2 using Viola&Jones features, trained in 64 stages

	Fear	Surprise	Sadness	Anger	Disgust	Happy
Fear	**52.94**	16.47	15.29	1.17	5.88	8.23
Surprise	12.14	**84.11**	1.86	0.93	0.93	0
Sadness	10.98	4.39	**71.42**	4.39	3.29	5.49
Anger	27.58	3.44	24.13	**24.13**	13.79	6.89
Disgust	10.16	0	11.86	5.08	**66.10**	6.77
Happy	4.58	0.91	0.91	0	0.91	**92.66**

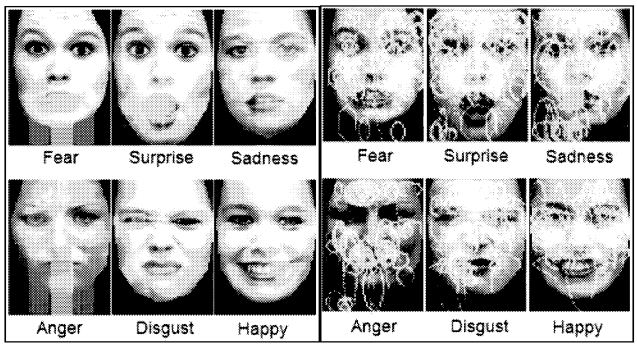


Figure 6: Graphical representation of the projection of Viola&Jones features (left) and Local Binary Patterns (right) selected by the optimal Adaboost.M2 classifiers on face samples representing basic facial expressions

Figure 6 depicts the sets of relevant LBP and Viola&Jones features selected by Adaboost.M2 classifiers for each class of facial expressions. The classifier which has the smallest misclassification rate for all facial expressions is called the optimal classifier and is obtained after 63 training steps. Because at each step one weak classifier is trained by using one feature only, the optimal classifier uses 63 features for each facial expression category. Table 1 shows the confusion matrix of Adaboost.M2 classifier (Freund and Schapire 1997) using LBP features as evaluated with leave-one-out cross validation. The overall test recognition rate of the classifier is 71.04% and the training mismatch rate is 0.22%.

Based on an image size of 60X80 pixels we have originally generated a set of 617.525 Viola&Jones basic features. The types (a),(b) and (c) each accounts with 27.65% to the total amount of features. We have prepared the final data set by uniformly sampling a third of the initial features. The procedure resulted to a set of 480 face samples, each being represented by a vector of 205.842 elements.

The optimal classification is obtained with a model trained in 45 stages. The overall test accuracy of this model is 72.29%, while the train mismatch rate is 0.40%. Table 2 shows the confusion matrix of this model.

CONCLUSIONS AND FUTURE WORK

In the current paper we have presented a method based on the original formulation of Active Appearance Models. As a novelty, we employed 3D face data for the recognition of user's affect. Further investigations are necessary to improve the model to fit the requirements of the environment and to find out to which extent 3D AAM performs better than 2D AAM. Another goal is to adapt the Active Appearance Model for stereo vision so as to increase the accuracy of 3D face analysis. Multimodal recognition of user's affect is also taken into consideration as a method which employs the fusion of audio-visual data.

REFERENCES

Cootes T.F.; C.J. Taylor; D.H. Cooper; and J. Graham. 1995. "Active shape models: their training and application." Computer Vision and Image Understanding, 61(1), 38–59.

Datcu D. and L.J.M. Rothkrantz. 2008. "Semantic audio-visual data fusion for automatic emotion recognition." Euromedia'2008, Porto, ISBN 978-9077381-38-0, Eurosis, Ghent, April, 58-65.

Datcu D. 2009. "Multimodal recognition of emotions", PhD. Thesis, Delft University of Technology, October.

Dornaika F. and J. Ahlberg. 2006. "Fitting 3D face models for tracking and active appearance model training Image and Vision Computing", 24(9), 1010-1024.

Edwards G.J.; C.J. Taylor; and T.F. Cootes. 1998. "Interpreting face images using active appearance models". In: FG '98: Proceedings of the 3rd. International Conference on Face & Gesture Recognition, IEEE Computer Society, Washington, DC, USA.

Ekman P. and W.V. Friesen. 1978. "Facial action coding system: investigator's guide." Consulting Psychologists Press, Palo Alto.

Freund Y. and R.E. Schapire. 1997. "A decision-theoretic generalization of online learning and an application to boosting." Journal of Computer and System Science, 55, 119–139.

Kanade T.; J.F. Cohn; and Y. Tian. 2000. "Comprehensive database for facial expression analysis." In Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG'00), 46–53.

Lefèvre S and J.M. Odobez. 2010. "View-Based Appearance Model Online Learning for 3D Deformable Face Tracking." In Proc. Int. Conf. on Computer Vision Theory and Applications (VISAPP), Angers, May.

Marcel S.; J. Keomany; and Y. Rodriguez. 2006. "Robust to illumination face localisation using shape models and local binary patterns." Technical report, IDIAP.

Ojala T.; M. Pietikäinen; and D. Harwood. 1996. “A comparative study of texture measures with classification based on feature distributions.” PR, 29(1), January, 51–59.

Ojala T.; M. Pietikäinen; and T. Mäenpää. 2002. “Multiresolution grayscale and rotation invariant texture classification with local binary patterns.” IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(7), 971–987.

Savran A.; N. Alyüz; H. Dibeklioğlu; O. Çeliktutan; B. Gökberk; B. Sankur; and L. Akarun. 2008. “Bosphorus Database for 3D Face Analysis”, The First COST 2101 Workshop on Biometrics and Identity Management (BIOID 2008), Roskilde University, Denmark, May.

Tong Y.; W. Liao; and Q. Ji. 2007. “Facial action unit recognition by exploiting their dynamic and semantic relationships.” IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(10), October, 1683–1699.

Viola P. and M. Jones. 2001. “Robust real-time object detection.” International Journal of Computer Vision.

Viola P. and M. J. Jones. 2004. “Robust real-time face detection.” IJCV:International Journal of Computer Vision, 57(2), May, 137-154.

A Semi-Automatic System for Posture Risk Assessment

Paulo J. S. Gonçalves[1,2] & Nuno O. G. Fernandes[1,3]
[1]Polytechnic Institute of Castelo Branco, School of Technology
Av. Empresário, 6000-767 Castelo Branco, Portugal.
[2]Technical University of Lisbon, IDMEC / IST
Av. Rovisco Pais, 1049-001 Lisboa, Portugal.
[3]CITEP, University of Minho
Campus de Azurém, 4800-058 Guimarães, Portugal.
email: paulo.goncalves@ipcb.pt , nogf@ipcb.pt

KEYWORDS
Posture Recognition, Vision System, Intelligent System

ABSTRACT

RULA (Rapid Upper Limb Assessment) is a survey method that assesses biomechanical and postural loading on the whole body with particular attention to the neck, trunk and upper limbs. In this paper we present a semi-automatic posture recognition system, to collect data and to evaluate human posture risk based on RULA. The proposed system uses two synchronized video cameras to capture images or posture samples of the operator performing its job. RULA scores are then automatically computed from these posture samples, based on semi-automatic and automatic image processing algorithms. These algorithms can extract visual features from the human upper limb, and obtain its model. The resulting scoring generates an action list which indicated the level of intervention required to reduce the risks of injury due to physical loading on the operator. The developed system is intended to be used as a tool for ergonomic investigations of workplaces.

INTRODUCTION

Micro, small and medium-sized enterprises (SMEs) play a central role in the European economy. In the enlarged European Union, SMEs represent 99% of all enterprises (Comission 2003). Despite this economical importance, workers in SMEs may be exposed to less favourable working conditions than the workers in large enterprises (O. Sorensen and Bach 2007). Cumulative trauma disorders (CTD) or work-related musculoskeletal disorders constitute an important occupational problem for this enterprises with rising costs of wage compensation and medical expenses, reduced productivity, and lower quality of life (Kee and Karwowski 2007).
Postural analysis can be a powerful technique for assessing work activities (Hignet and McAtamney 2000). There have been several research techniques proposed in the literature to assess work activities and the job risk level, e.g. OWAS - Ovako Working posture Analysis System (O. Karhu and Kuorinka 1977), REBA - Rapid Entire Body Assessment (Hignet and McAtamney 2000); LUBA - loading on the upper body assessment (Kee and Karwowski 2007), (M.K. Chung and Kim 2002) and RULA (McAtamney and Corlett 1993), between others.
RULA is a validated method that assesses biomechanical and postural loading on the upper limbs. It was developed to identify work posture or risk factors that deserve further attention. Positions of individual body segments are observed and the more there is a deviation from the neutral posture the higher will be the score of each body part. Additional weights are given to the postures by loads handled and repetitive muscular activity. Scores for the upper arm, hand/wrist, trunk and lower extremity are combined in a final RULA score that quantifies the potential risk due to physical loading on the operator. An action list is then generated which indicates the level of intervention required to reduce the risk factors. Several applications of RULA in practical situations have been described in the literature, e.g. (S. Oates and Hedge 1998), (M. Massaccesi and Greco 2003), (E.C. Lee and Dennerlein. 2005) amongst others.
It is a common practice in posture risk evaluations to use a pre-defined posture categorization system for the relative positions of the segmented parts of the human body (S. Bao and Silversteina 2007) and the analyst only needs to decide on which category a joint angle is. However, this approach tends to be time consuming to be useful in the analysis of complex tasks, particularly if these tasks are dynamic or varied and the work cycle is long. In this case the analyst usually takes an assessment at periodic time intervals. Another limitation, accordingly to (S. Bao and Silversteina 2007), is the classification bias that may occur if the analyst has some advanced knowledge about the working or health condition of the worker observed. For example the posture may be placed in the higher risk posture category if the analyst knows the work condition is poor. In order to overcome these limitations we propose a computerised semi-automatic posture recognition system, which uses

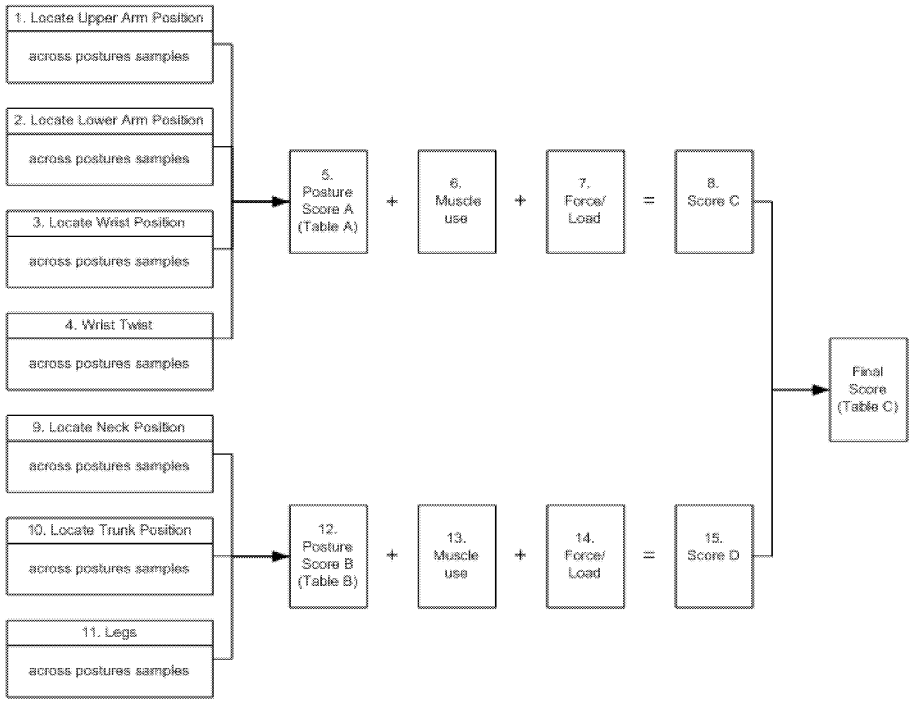


Figure 1: RULA methodology for assessing the risk of upper limb disorders.

two synchronized cameras to capture the images of the operator performing its job from different angles. A time-based approach is adopted to collect posture samples at random time intervals and the RULA scores are then computed for these posture samples. The remainder of the paper is organized as follows. The next sections present the RULA procedure adopted in the proposed automatic posture recognition system and detail the semi-automatic measuring system. Next the experimental results obtained from the developed system are presented. Concluding remarks and directions for future research work are put forward in the final section of the paper.

THE RULA PROCEDURE ADOPTED IN THE SEMI-AUTOMATIC SYSTEM

The semi-automatic takes posture samples using the two synchronized video cameras at regular time intervals. This is particular useful when the work cycle is long or the postures are varied. Because RULA assessments are taken at set time intervals over the working period the proportion of time spent in the various postures (posture frequency) is evaluated. For each posture sample the position of the upper arm, lower arm and wrist are assessed and a posture score A is obtained. This is done by, first assessing the ranges of movement (see following section) and then by multiplying the corresponding risk index (based on the RULA method) by the posture frequency. Posture score A is then obtained from the Upper limb posture score table of the RULA method, in (Gonçalves and Fernandes 2010), using the previous obtained risk indexes for the upper arm, lower arm, wrist and wrist twist. Figure 1 describes the general methodology for assessing the risk of upper limb disorders. In the same way the position of the neck, trunk and legs are assessed to obtain a posture score B using the Neck, trunk and legs posture score table of the RULA method, in (Gonçalves and Fernandes 2010). This gives an initial estimate of the potential risk of the task. Force and frequency are then taken into account and must be introduced by the analyst. This allows modifying the initial estimate to take into account the effects of the loadings imposed. Having incorporated the values of these two factors these are now summed as indicated in Figure 1 to obtain posture scores C and D. From these two scores, a final score is obtained using Table C in appendix. Based on the RULA final score, four action levels are suggested by (McAtamney and Corlett 1993), which give an estimation of the potential risk of musculoskeletal symptoms arise and indicate the level of intervention required to reduce that risk. The greater the score, the higher the risk:

- *Score 1 or 2*: acceptable posture,
- *Score 3 or 4*: further investigation and changes may be needed,
- *Score 5 or 6*: investigation and changes are required soon,
- *Score 7*: investigation and changes are required immediately.

Since the human body is a complex and adaptive system, the action level list must be seen as a guide for further action.

RISK ASSESSMENT

Semi-automatic Risk Assessment

This section presents the approach proposed in (Gonçalves and Fernandes 2010), for a semi-automatic risk assessment system. It establishes risk indexes for body segments, based on the range of movement. An example is provided for upper arm position.
Following the Rula Method (McAtamney and Corlett 1993), the system establishes the risk index for the upper arm position which identifies three basic features from the acquired images, namely: the degree of extension/ flexion of the upper arm; upper arm abduction and the shoulder position. Information about the arm support, wrist position, muscle use and force/load applied is considered to be introduced by the user, i.e. the analyst.
Semi-automatic algorithms are used to select and measure the degree of extension/ flexion of the upper arm; upper arm abduction and the shoulder position. Posture samples are obtained using two synchronized video cameras located at fixed positions around the worker, allowing two different viewpoints: frontal and lateral.
The degree of extension/flexion of the upper arm is measured, in the lateral view, from the angle between the line segment representing the upper arm and the line segment perpendicular to the ground floor, i.e., the vertical line. The first line segment is identified by the user, i.e. the analyst, when selecting two points in the capture image. These points must be selected in the acromion (shoulder) and in the palpable extremity of radius (elbow). When placing both cameras in a plane parallel to the ground floor, the second line segment is immediately defined by each one of the columns of the image matrix. The angle is measured in the intersection between the two previous defined line segments, e.g., the two corresponding vectors $\overrightarrow{u}$ and $\overrightarrow{v}$, using:

$$\theta_e = cos^{-1}\left(\frac{\overrightarrow{u}\cdot\overrightarrow{v}}{\|\overrightarrow{u}\|\cdot\|\overrightarrow{v}\|}\right) \quad (1)$$

According to (McAtamney and Corlett 1993), the relation between the feature θ_e and extension/flexion is: if $\theta_e > 0^o$ then the upper arm is in flexion, else if $\theta_e < 0^o$ then the upper arm is in extension. Then the algorithm is: if $\theta_e < -20^o$ then risk index is 1, else if $-20^o < \theta_e < 20^o$ then risk index is 1, else if $20^o < \theta_e < 45^o$ then risk index is 2, else if $45^o < \theta_e < 90^o$ then risk index is 3, else if $\theta_e > 90^o$ then risk index is 4.
The upper arm abduction is measured, in the frontal view, from the angle between the line segment representing the upper arm and the line segment perpendicular to the ground floor. The procedure is precisely the same as for the extension/flexion feature, giving the feature θ_a. According to (McAtamney and Corlett 1993), if $\theta_a > 0^o$ then the risk index is 1; else the risk index is 0. The shoulders relative position is measured, in the frontal view, from the angle between the line between the two shoulders and the horizontal. The first line segment is identified by the user when selecting two points in the capture image. Points must be placed in the acromion bones. When placing both cameras in a plane parallel to the ground floor, the second line segment is immediately defined by each one of the lines of the image matrix. The angle is measured in the intersection between the two previous defined line segments, e.g., the two corresponding vectors $\overrightarrow{s}$ and $\overrightarrow{h}$, using:

$$\theta_s = cos^{-1}\left(\frac{\overrightarrow{s}\cdot\overrightarrow{h}}{\|\overrightarrow{s}\|\cdot\|\overrightarrow{h}\|}\right) \quad (2)$$

According to (McAtamney and Corlett 1993), if $\theta_s > 0^o$ then the risk index is 1, else the risk index is 0.

Automatic Posture Recognition

The method followed in this paper to automatically recognize workers posture, for the upper-body case, is presented in (Ferrari et al. 2008). In the upper-body case there are six parts to be identify: head, trunk and left and right arms and forearms. The pose recognition method estimates the parameters of a 2D articulated model, i.e., the (x, y) position location of each of the six body parts, its orientation and scale.
First, is compulsory to detect humans in the image. This goal can be achieved by matching a 2D articulated model to the whole captured image, but is time consuming. In (Ferrari et al. 2008) is proposed a method to detect the worker that is based in a generic upper-body detector (Dalal and Triggs 2005), i.e., the called weak model. This detection is used in the next steps of the algorithm to reduce the search space of the 2D articulated model, to a region of interest (ROI).
The second step of the method removes part of the background clutter, using the Grabcut segmentation (C. Rother and Blake 2004), highlighting the foreground. This step allows the third, that outputs a first pose recognition using the image parsing technique proposed in (Ramanan 2007).
In the last step of the method a fine detection of the articulated model is performed using a spatio-temporal parsing, as described in (Ferrari et al. 2008). This approach uses (ii) appearance models integrated from multiple frames where the system is confident about the estimated pose; and (ii) infers over a joint spatio-temporal model of pose, capturing both kinematic constraints within a frame, and temporal continuity constraints between frames.
The system was trained with the Weizmann dataset (L. Gorelick and Basri 2007) using a linear SVM (Cortes

Figure 3: Upper arm position.

Figure 4: Shoulder is raised?

and Vapnik 1995), 88% accuracy.

EXPERIMENTAL RESULTS

Semi-automatic Risk Assessment

In this section are presented the experimental results obtained for the upper arm and wrist position of an office worker. In Figure 2 is depicted the graphic user interface (GUI) developed, using MatLab2009b. The office worker is shown from two camera angles at a randomly pre-selected video frame: front position (camera 1) and lateral position (camera 2). The software used for the risk calculus allows the interaction with the user (analyst), to define the visual feature points, to enter the arm support, wrist position, muscle use, force/load applied and to visualize the results.

Figure 3 illustrates the measurement the upper arm extension/ flexion angle θ_e . Since $\theta_e \approx 45^o$, in flexion, the associated risk factor is +3. Figure 4 shows the shoulders relative position by measuring the angle between the line of the shoulders and the horizontal (neutral position of the shoulders). Since $\theta_s \approx 10^o$, the adjustment is +1. Figure 5 shows the situation where the upper arm abduction angle is determined. Since $\theta_a \approx 45^o$, the adjustment is +1. Figure 6 shows the

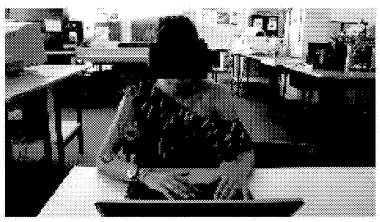


Figure 5: Upper arm is abducted?

Figure 6: Arm angle.

situation where the arm angle is determined. Since the angle $\approx 131^o$, the adjustment is +2. The step 5 of the algorithm, see Figure 1, gives the score 6. For steps 6 and 7, the score is 0, which gives 6 has the final score A, as depicted in Figure 2. This indicates that investigation and changes are required soon (McAtamney and Corlett 1993)

Applying the above algorithms requires that the user introduces two points in the captured frames. These are related with the acromion position and/or the palpable extremity of radius. This is the main source of error in the measurement process, since the exact position of the points in the captured frames cannot be exactly identified, which might be hidden by, for example, the cloth. So the points introduced by the user in the captured frames must be seen as an approximation to the exact position. However this has minor implication for assessing risk indexes, since the RULA method is based on the range of movements.

Automatic Posture Recognition

This section presents the first results for full automation of the upper limb visual features extraction. This is accomplished using the algorithms presented in the previous section. In Figure 7 are presented the line segments that define the human upper limb: one for the neck, one for the trunk and one for each left and right arm and forearm. As an example for RULA visual feature extraction, the angle between the arm and the forearm is calculated using equation (1). For the left and right parts of the body, the angles are approximately 151 and 139 degrees, respectively.

CONCLUSIONS AND FUTURE WORK

The paper presents a semi-automatic system for posture risk assessment. The method was specially designed to be used in SMEs, where workers may be exposed to less favorable working conditions than the workers in large enterprises.

The proposed system is low cost and portable, due to the use of webcams. This kind of cameras were used because angles calculation accuracy is not required to

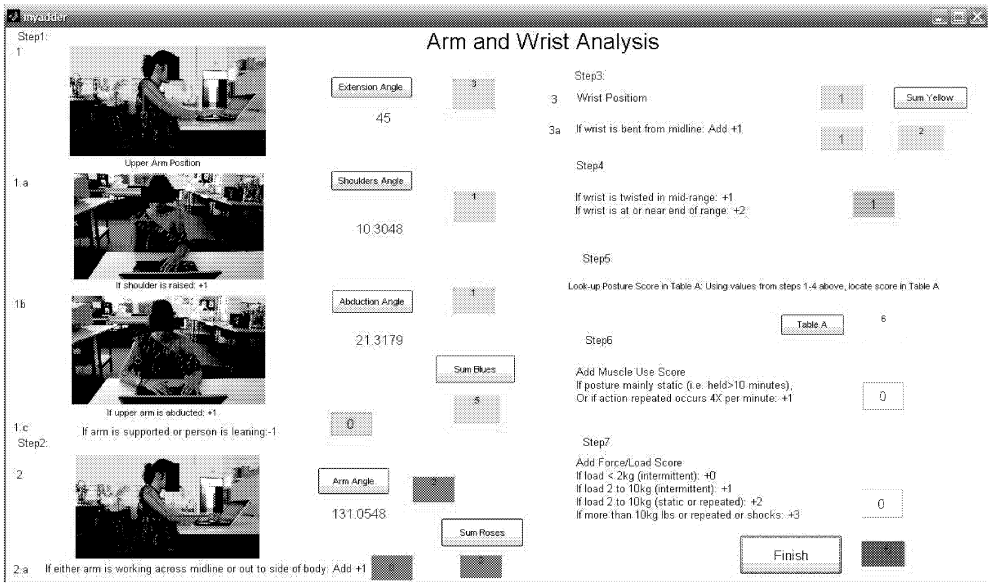


Figure 2: The Semi-Automatic System for Posture Risk Assessment.

Figure 7: Automatic recognition of arms, head and trunk.

be extremely high, since the RULA method relies on a range of values and the scores obtained are indicators for future actions.

With the proposed system the posture risk assessment is performed very quickly when compared with the actual procedure, where the analyst performs risk analysis only based on observations of the work cycle. The system fulfils the need of providing a method for screening a large number of operators quickly and minimizing the classification bias that may be introduced by analysts.

The proposed semi-automatic system for posture risk assessment was applied with success for accessing the risk indexes of the upper arm position, using two approaches. The first approach requires that the analyst introduces in the captured frames reference points for the parts of the human body to measure. Since this procedure can be seen as time consuming the second approach proposed is focused in decreasing the intervention of user by automatically determination of the RULA visual features. Examples were given that validates the approaches.

Future work will be focused in increasing the accuracy and decreasing the user intervention on the semi-automatic system for posture risk assessment.

REFERENCES

C. Rother V.K. and Blake A., 2004. *GrabCut - Interactive Foreground Extraction using Iterated Graph Cuts.* In *ACM Transactions on Graphics (SIGGRAPH).*

Comission E., 2003. *SME.* `http://ec.europa.eu/enterprise/policies/sme/facts-figures-analysis/sme-definition/index_en.htm`.

Cortes C. and Vapnik V., 1995. *Support-Vector Networks. Machine Learning*, 20, 273–297.

Dalal N. and Triggs B., 2005. *Histograms of oriented gradients for human detection.* In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.*

E.C. Lee A. Rafiq R.M.R.A. and Dennerlein. J., 2005. *Ergonomics and human factors in endoscopic surgery: a comparison of manual vs telerobotic simulation systems. Surgical Endoscopy*, 19, no. 4, 1064–1070.

Ferrari V.; Marin-Jimenez M.; and Zisserman A., 2008. *Progressive Search Space Reduction for Human Pose Estimation.* In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.*

Gonçalves P. and Fernandes N., 2010. *RULAMATIC - A Semi-Automatic Posture Recognition System for Rula Risk Evaluation Method. Romanian Review Precision Mechanics, Optics and Mechatronics*, 37, 119–124.

Hignet S. and McAtamney, 2000. *Rapid Entire Body Assessment (REBA). Applied Ergonomics*, 31, 201–205.

Kee D. and Karwowski W., 2007. *A Comparison of Three Observational Techniques for Assessing Postural Loads in Industry. International Journal of Occupational Safety and Ergonomics*, 13, no. 1, 3–14.

L. Gorelick M. Blank E.S.M.I. and Basri R., 2007. *Actions as Space-Time Shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29, no. 12, 2247–2253.

M. Massaccesi A. Pagnottaa A.S.M.M.C.M. and Greco F., 2003. *Investigation of work-related disorders in truck drivers using RULA method. Applied Ergonomics*, 34, no. 4, 303–307.

McAtamney L. and Corlett E., 1993. *RULA: a survey method for the investigation of work-related upper limb disorders. Applied Ergonomics*, 24, no. 2, 91–99.

M.K. Chung I.S. Lee D.K. and Kim S., 2002. *A postural workload evaluation system based on a macro-postural classification. Human Factors and Ergonomics in Manufacturing*, 12, 267–277.

O. Karhu P.K. and Kuorinka I., 1977. *Correcting working postures in industry: a practical method for analysis. Applied Ergonomics*, 8, 199–201.

O. Sorensen P.H. and Bach E., 2007. *Working in small enterprises - Is there a special risk? Safety Science*, 45, no. 10, 1044–1059.

Ramanan D., 2007. *Learning to parse images of articulated bodies.* In B. Schölkopf; J. Platt; and T. Hoffman (Eds.), *Advances in Neural Information Processing Systems 19*, MIT Press, Cambridge, MA. 1129–1136.

S. Bao N. Howarda P.S. and Silversteina B., 2007. *Two posture analysis approaches and their application in a modified Rapid Upper Limb Assessment evaluation Ergonomics. Ergonomics*, 50, no. 12, 2118–2136.

S. Oates G.E. and Hedge A., 1998. *An Anthropometric and Postural Risk Assessment of Children's School Computer Work Environments. Computers in the Schools*, 13, no. 3.

3D SURFACE REGISTRATION OF ULTRASOUND TO CT BONE IMAGES

Pedro M. B. Torres[1], Paulo J. S. Gonçalves[1,2], Jorge M. M. Martins[2]
[1]Polytechnic Institute of Castelo Branco, School of Technology,
Av. Empresário, 6000-767 Castelo Branco, Portugal
[2]Technical University of Lisbon, IDMEC / IST,
Av. Rovisco Pais, 1049-001 Lisboa, Portugal
email: {pedrotorres|paulo.goncalves}@ipcb.pt, martins@dem.ist.utl.pt

KEYWORDS
Image-guided surgery, Surface-based registration

ABSTRACT

This paper presents an approach for the 3D surface-based registration using the ICP algorithm. A rigid registration of B-Mode Ultrasound (US) images to Computer Tomography (CT) images of a cow femur is presented. This registration is suited to help robot navigation in orthopedic surgery and can also be used in the preoperative scenario to surgical planning. The paper presents results that demonstrate the validity of the method.

INTRODUCTION

In recent years robots are starting to work in the operating room, to help surgeons in surgical procedures (e.g. ROBODOC, DaVinci robot, ACROBOT, etc). Orthopedic surgery is a very delicate surgery, where problems for patients can appear, when positioning errors occur. In hip resurfacing surgery, the insertion of a spike in the femur requires high accuracy and precision in its position and orientation. To minimize after surgery problems our research group is developing a robot to help surgeons. The development of a robotic arm, allows to reduce the complexity of the surgical procedures reducing the efforts of the surgeon, surgery time and the effects of invasive surgery. To control the robot is necessary to develop a image-guidance system. In recent years,(Ecker and Timo M. Tannast 2007),(Amiot and Poulin 2004), (Barger et al. 1998), (Nabeyama et al. 2004) several solutions were proposed, based on the use of fiducial markers to obtain the position and orientation of the bone relatively the structure of robot. Ultrasound (US) images of bones can be used to overcome the bone incisions to attach the fiducial markers in orthopedic surgery, and to extract the position and orientation of the bone for robotic navigation during the surgical procedure. However, there are a number of image issues associated with US: speckle noise, saturation of the reflected echo at the bone-tissue boundary and variation of the speed of sound in different tissues, can all make precise location of the bone surface difficult (Penney et al. 2006). The main idea of the navigation system, is to acquire CT images of the femur in a preoperative scenario, perform 3D reconstruction of these images, acquire US images of the femur in the intra-operative scenario, perform 3D reconstruction of these images and finally register the two surfaces obtained. In this work, we use a cow femur bone, as a study model. Bone CT images, were acquired with spacing of 0.75 mm, and US images were acquired in B-Mode with the help of a tub of water with spacing of 3 mm, to enable the acquisition of images with minimal dispersion. The image processing and surface registration is performed in $MATLAB$, but it was also used $MeVisLab$ and $MeshLab$ for reconstruction and visualization.
This paper is organized as follows. In the next section we describe the process of image registration based on 3D surfaces and also the Iterative Closest Point algorithm. The next section describes the experimental setup and presents the obtained results. Finally, conclusions and future work are presented.

3D SURFACE BONE REGISTRATION

Several methods have been developed in recent years to perform the registration of bone ultrasound images, obtained in the intra-operative scenario, to the 3D bone model, obtained from CT images (Winter and Brendel 2005).

The approaches presented by the research community are nowadays a active field of research, due to the need of a high precision system. In fact, the following approaches can perform US images to 3D bone model registration on several parts of the human body but still need to improve its overall accuracy. The approaches tackle the Spine (Brendel et al. 2002), the shoulder (Tyryshkin et al. 2007) and the nose (Descoteaux et al. 2006). None of them tackled the femur.
The method proposed by (Brendel et al. 2002) estimates the translation and rotation parameters between images

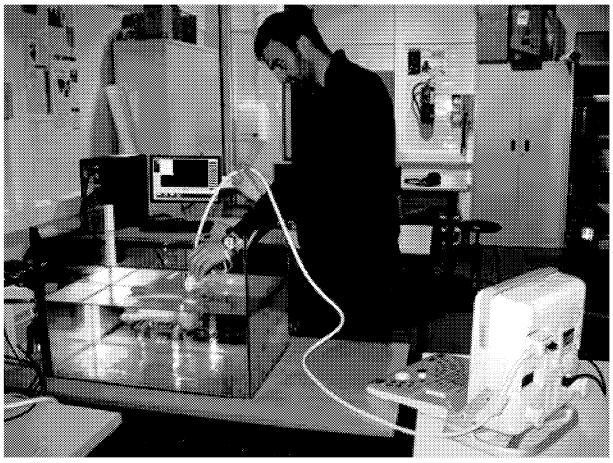


Figure 1: Experimental Setup.

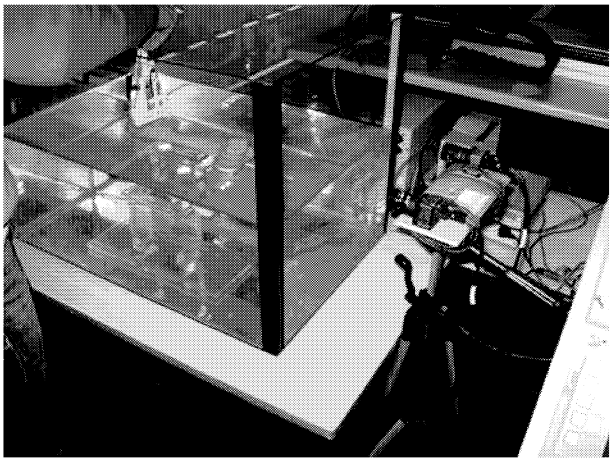


Figure 2: US probe with markers and Stereo Vision System.

and is based in the mean value of the pixel grey values. The nearest neighbor, k-NN (Dasarathy 1991), approach estimates the location of the correspondent grey area between images. Finally the Levenberg-Marquart optimization method (Levenberg 1944, Marquardt 1963) is used to estimate the rotation and translation parameters. The sum of the pixel bright grey values of the US image, i.e. the surface of the bone, are used by (Brendel et al. 2008) to calculate the correspondences between images.
(Beek et al. 2008) uses 3D points, obtained from an 3D US machine and from the classical TC, to apply the Iterative Closest Point (ICP) algorithm (Besl and McKay 1992), to obtain the translation and rotation parameters of the transformation. The mutual information approach is used by (Chen and Abolmaesumia 2005) to register intraoperative US images to preoperative US images. The preoperative images are also used to construct the 3D volume of the bone. When the 2D intraoperative US image is registered to the preoperative US image, the 3D position od the bone in the operating room can be easily known.
The above presented methods, apart from digital image processing to obtain the bone contours, all rely on the registration of the 3D contour points obtained from US to the ones obtained from CT images (preoperative 3D bone model). The most commonly method used is the ICP. Recently (Myronenko and Song 2010) have proposed a new method, that relies on a probabilistic framework to perform the registration.

Iterative Closest Point Method

The Iterative Closest Point (ICP) method, presented by (Besl and McKay 1992) is the standard method used to perform registration between two set of 3D points. It transforms two sets of points to a common coordinate frame. If the exact correspondences of the two data set could be known, then the exact translation t and rotation R can be found. The main issue of the method is then to find the corresponding points between the two data sets, $Y = (y_1, ..., y_M)^T$ and $X = (x_1, ..., x_N)^T$.
The assumption in the ICP method is that the closest points between the data sets correspond to each other, and are used to compute the best transformation, rotation and translation, between them. The original method have been extended to line segment sets, implicit curves, parametric curves, triangle sets, implicit surface and parametric surfaces.
To obtain the closest point of Y to a point in X, the Euclidean distance is applied:

$$d(Y, X) = \sqrt{(X_x - Y_x)^2 + (X_y - Y_y)^2 + (X_z - Y_z)^2} \quad (1)$$

When all points of the data set Y are associated to the point in X the transformation is estimated by minimizing a mean square cost function:

$$E_{ICP} = \sum_i \|R \cdot x_i + t - y_i\|^2 \quad (2)$$

From the obtained parameters, the points in the X data set are transformed and the error between them and the ones in Y calculated. If the error is above a predefined threshold then the points must be re-associated and the previous steps again performed until the error is below the threshold. Figure 3 helps us understand the alignment of sets of points of two lines.

EXPERIMENTAL EVALUATION

Experimental Setup

For obtaining the CT images, a Siemens commercial machine was used with 0.75 mm between slices. The US images were acquired through the experimental apparatus shown in figure 1, which consist of a tub with water (to place the femur of cow and its support), the $ALOKA$

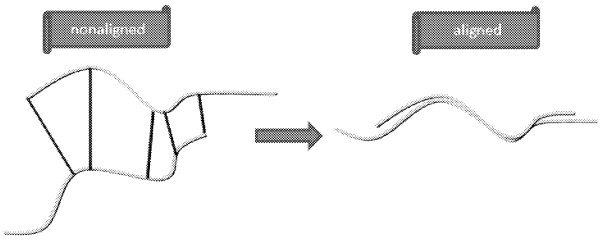


Figure 3: Example of alignment for local minimum.

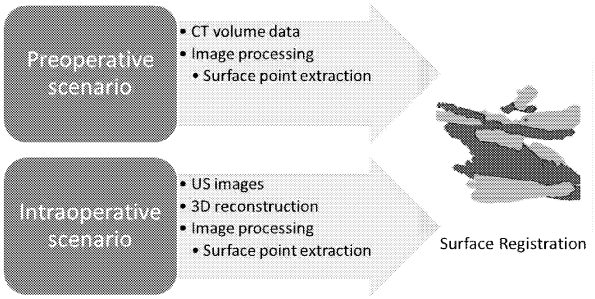


Figure 4: Block diagram of the structure followed.

prosound 2 US machine, a PC equipped with a standard video capture board and a stereo video system(Morgado et al. 2009) to determine the 3D coordinates of the markers in the ultrasound probe, depicted in figure 2. The steps performed are presented in the diagram of figure 4. CT images are acquired in the preoperative scenario and are processed to obtain de 3D surface point reconstruction. In intraoperative scenario, US images of bone are acquired using the *ALOKA* machine and the standard video capture board. The images are processed and 3D surface are obtained by 3D reconstruction using triangulation. Registration of the two 3D surfaces is obtained by ICP algorithm, described in the previous section.

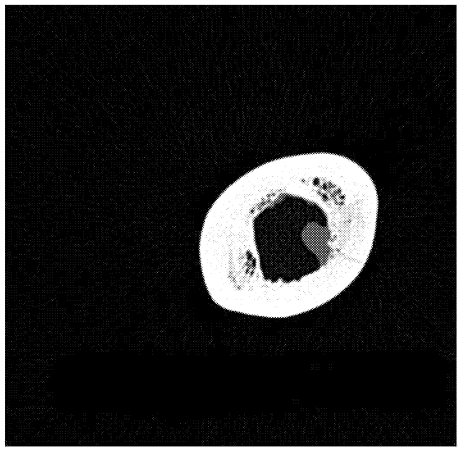


Figure 5: A Bone CT slice.

Figure 6: MeVisLab 3D bone Model.

CT Image Processing

In the acquisition of CT images, 360 images were saved by dataset, since the size of the bone is 270 *mm* and the spacing between slices is 0.75 *mm*, figure 5 shows a bone CT slice. Femur 3D model, was obtained by reconstruction of the 360 slices, using the Open Source Software *MeVisLab*, figure 6. To perform registration between CT data and US data, was considered only the top layer of the bone as region of interest (ROI), since in the US images is only visible the bone top, because US waves are reflected in the bone, figure 7. Through surface triangulations, is obtained the 3D reconstruction of the interest surface, which will be used in the registration process with the surface obtained from US images. Figure 8 shows the 3D reconstruction obtained with the Open Source Software *MeshLab*.

US Image Processing

Ultrasound B-Mode images were acquired with probe of 5 MHz. To perform surface registration, 74 images were acquired using continuous scanning, with 3 *mm* spacing between them. Figure 9 shows an US image on the central area of the femur. As the US images are always with noise, it was necessary to clean the speckle noise and segment the bone region of interest. This process was performed by De-speckling algorithms (Seabra and Sanches 2010) and morphological operators to eliminate isolated points in the images. Figure 10 shows the result of US image segmentation of the femur bone. After bone segmentation, a 3D point cloud obtained from the 74 images, is extracted using the stereo system. The later calculates the 3D position of the US probe and after

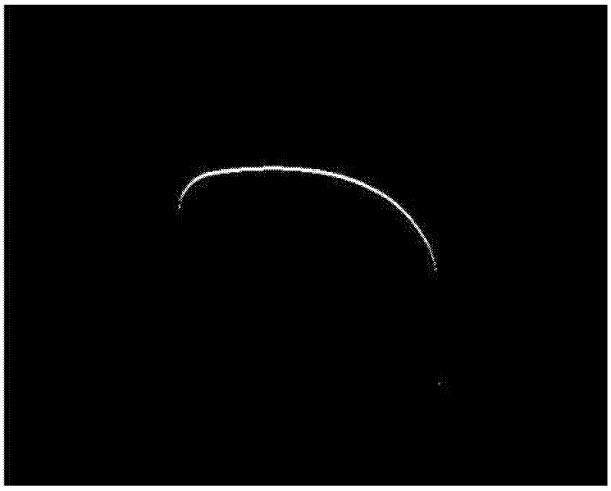


Figure 7: A ROI Bone CT slice.

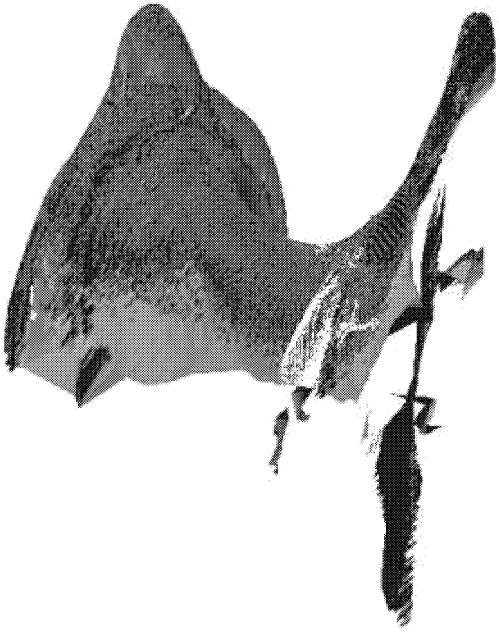


Figure 8: 3D reconstruction of the CT ROI.

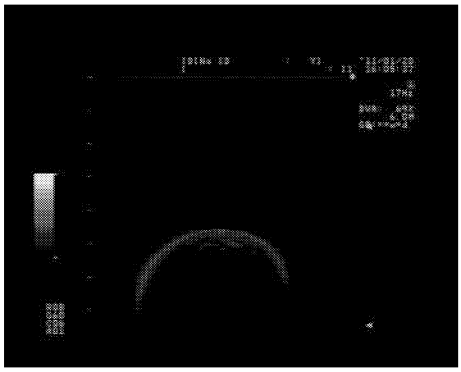


Figure 9: US image on the central area of the femur.

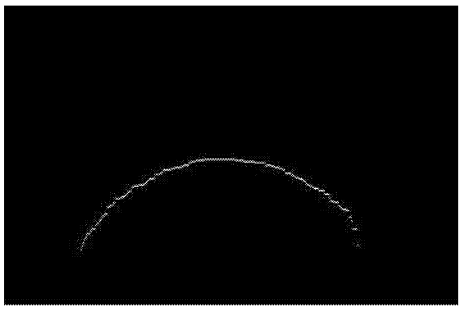


Figure 10: US image segmentation of the femur bone.

calibration the 3D position of each point of the bone contour for each one of the 74 images. The point cloud allows 3D surface reconstruction, by triangulation, as illustrated in figure 11.

Surface-based Registration Results

In this section are presented the Surface registration results obtained using the surfaces presented in figures 8 and 11. To perform registration, the ICP algorithm described in the previous section was used. A rigid registration is performed to match the two surfaces. From the output of the ICP method (the registered 3D US surface) the root mean square errors, RMSE,(3) was obtained related to the 3D surface of the CT data. A Intel Core 2 Duo, 2.27 GHz computer was used to perform the processing in *MATLAB* software. The algorithm converges after 29 seconds with the errors obtained before and after registration, presented in table 1. Despite the error have a high value, in this first phase of work is acceptable, since this is a global error, estimated from all points that make up the US and CT surfaces. Is important to note, as the algorithm is able to relate the two surfaces and this can be seen in figures 12 and 13, which illustrate the data before and after registration.

$$RMSE = \sqrt{\frac{1}{n}\sum(\widehat{y}_i - y_i)^2} \tag{3}$$

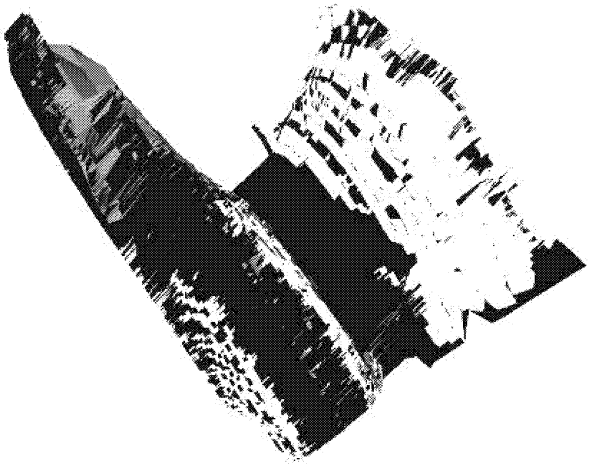


Figure 11: 3D reconstruction of the US surface.

Table 1: Registration errors.

ICP Method	RMSE [mm]
Before	83.6850
After	52.8625

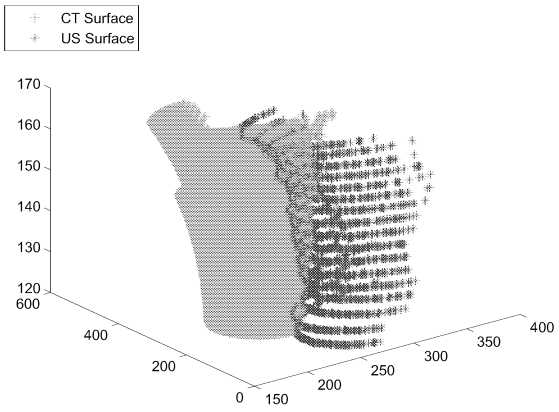


Figure 12: Surfaces before registration

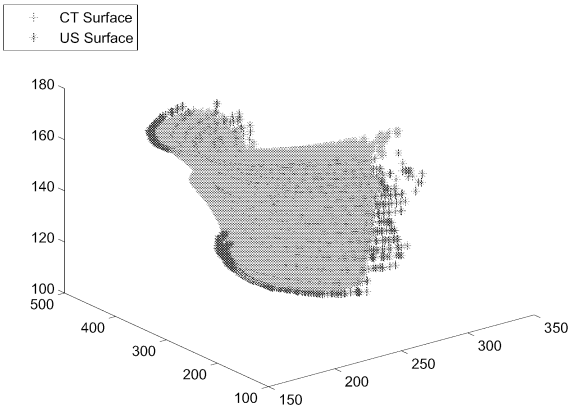


Figure 13: Surfaces after registration

CONCLUSIONS

This paper presents the registration of a 3D surface obtained from femur bone US images to its 3D model, obtained from CT images. The 3D surface, obtained from US images in the intraoperative scenario is then registered to the preoperative scenario. The ICP algorithm, commonly used for point cloud registration, was used for registration of the two 3D surfaces. The 3D surface corresponding to the US image data, is highly irregular (figure 11), not only by the spacing between slices (far superior to CT data) but also by the difficulty in extracting the surface points of the bone in US images, due to noise. However, the ICP method was able to register the two surfaces with accuracy. The results obtained validate the approach for the femur application. To achieve better accuracy, the data acquisition system must be enhanced by the well known Polaris Spectra state-of-the-art system. The bone contour extraction procedure will also be enhanced in order to speed up the system and reduce the errors.

Acknowledgements
The authors would like to thank: the Portuguese Science Foundation, FCT, for the funding to IDMEC through LAETA; the FCT project: PTDC/EME-CRO/099333/2008.

REFERENCES

Amiot L. and Poulin F., 2004. *Computed tomography-based navigation for hip, knee, and spine surgery. Clin Orthop Relat Res*, 421, 77–86.

Barger W.; Bauer A.; and Borner M., 1998. *Primary and revision total hip replacement using the ROBODOC' system. Clin Orthop Relat Res*, 354, 82–91.

Beek M.; Abolmaesumi P.; Luenam S.; Ellis R.; Sellens R.; and Pichora D., 2008. *Validation of a new surgical procedure for percutaneous scaphoid fixation using intra-operative ultrasound. Medical Image Analysis*, 12 (2), 152–162.

Besl P.J. and McKay N.D., 1992. *A Method for Registration of 3-D Shapes. Pattern Analysis and Machine Intelligence*, 14(2), 239 – 256.

Brendel B.; Winter S.; Pechlivanis I.; Schmieder K.; and C. I., 2008. *Registration of CT and Intraoperative 3-D Ultrasound Images of the Spine Using Evolutionary and Gradient-Based Methods. IEEE Transactions on Evolutionary Computation*, 12 (3), 284–296.

Brendel B.; Winter S.; Rick A.; Stockheim M.; and Ermert H., 2002. *Registration of 3D CT and Ultrasound Datasets of the Spine using Bone Structures. Computer Aided Surgery*, 7 (3), 146–155.

Chen T. and Abolmaesumia P., 2005. *A Mutual Information Based Registration Algorithm for Ultrasound-guided Computer-assisted Orthopaedic Surgery. Medical Imaging 2005: Visualization, Image-Guided Procedures*, 5744, 111–122.

Dasarathy B.V., 1991. *Nearest Neighbor (NN) Norms: NN Pattern Classification Techniquesy.* IEEE Computer Society. ISBN 0-8186-8930-7.

Descoteaux M.; Audette M.; Chinzei K.; and Siddiqi K., 2006. *Bone enhancement filtering: application to sinus bone segmentation and simulation of pituitary surgery. Computer aided surgery : official journal of the International Society for Computer Aided Surgery*, 11, 247–255.

Ecker and Timo M. Tannast Moritz Murphy S.B., 2007. *Computed tomography-based surgical navigation for hip arthroplasty. Clinical orthopaedics and related research*, 465, 100–105.

Levenberg K., 1944. *A Method for the Solution of Certain Non-Linear Problems in Least Squares. The Quarterly of Applied Mathematic*, 2, 164 – 168.

Marquardt D., 1963. *An Algorithm for Least-Squares Estimation of Nonlinear Parameters. SIAM Journal on Applied Mathematics*, 11, 431 – 441.

Morgado P.; Caldas Pinto J.; Martins J.M.M.; and Gonçalves P., 2009. *Cooperative Eye-in-hand/Stereo Eye-to-hand Visual Servoing.* In *Proc. of RecPad 2009 - 15th Portuguese Conference in Pattern Recognition.* Aveiro, Portugal.

Myronenko A. and Song X., 2010. *Point Set Registration: Coherent Point Drift. IEEE Transactions on Pattern Analysis and Machine Intelligence*, 99, no. PrePrints. ISSN 0162-8828. doi:http://doi.ieeecomputersociety.org/10.1109/TPAMI.2010.46.

Nabeyama R.; Matsuda S.; Miura H.; Mawatarim T.; Kawano T.; and Iwamoto Y., 2004. *The accuracy of image-guided knee replacement based on computed tomography. J Bone Joint Surg*, 86-B, 366–371.

Penney G.P.; Barratt D.C.; Chan. C.S.K.; Slomczykowski M.; Carter T.J.; Edwards P.J.; and Hawkes D.J., 2006. *Cadaver validation of intensity-based ultrasound to CT registration. Medical image analysis*, 10, 385–395.

Seabra J.C. and Sanches J., 2010. *ON ESTIMATING DE-SPECKLED AND SPECKLE COMPONENTS FROM B-MODE ULTRASOUND IMAGES. IEEE International Symposium on Biomedical Imaging (ISBI).*

Tyryshkin K.; Mousavi P.; Beek M.; Ellis R.; Pichora D.; and Abolmaesumi P., 2007. *A navigation system for shoulder arthroscopic surgery. Proceedings of the Institut of Mechanical Engineers Part H, Journal of engineering in medicine*, 6 (1), 221–19.

Winter S. and Brendel B. Igel C., 2005. *Registration of bone structures in 3D ultrasound and CT data: Comparison of different optimization strategies. International Congress Series*, 242–247.

AUTHOR LISTING

AUTHOR LISTING