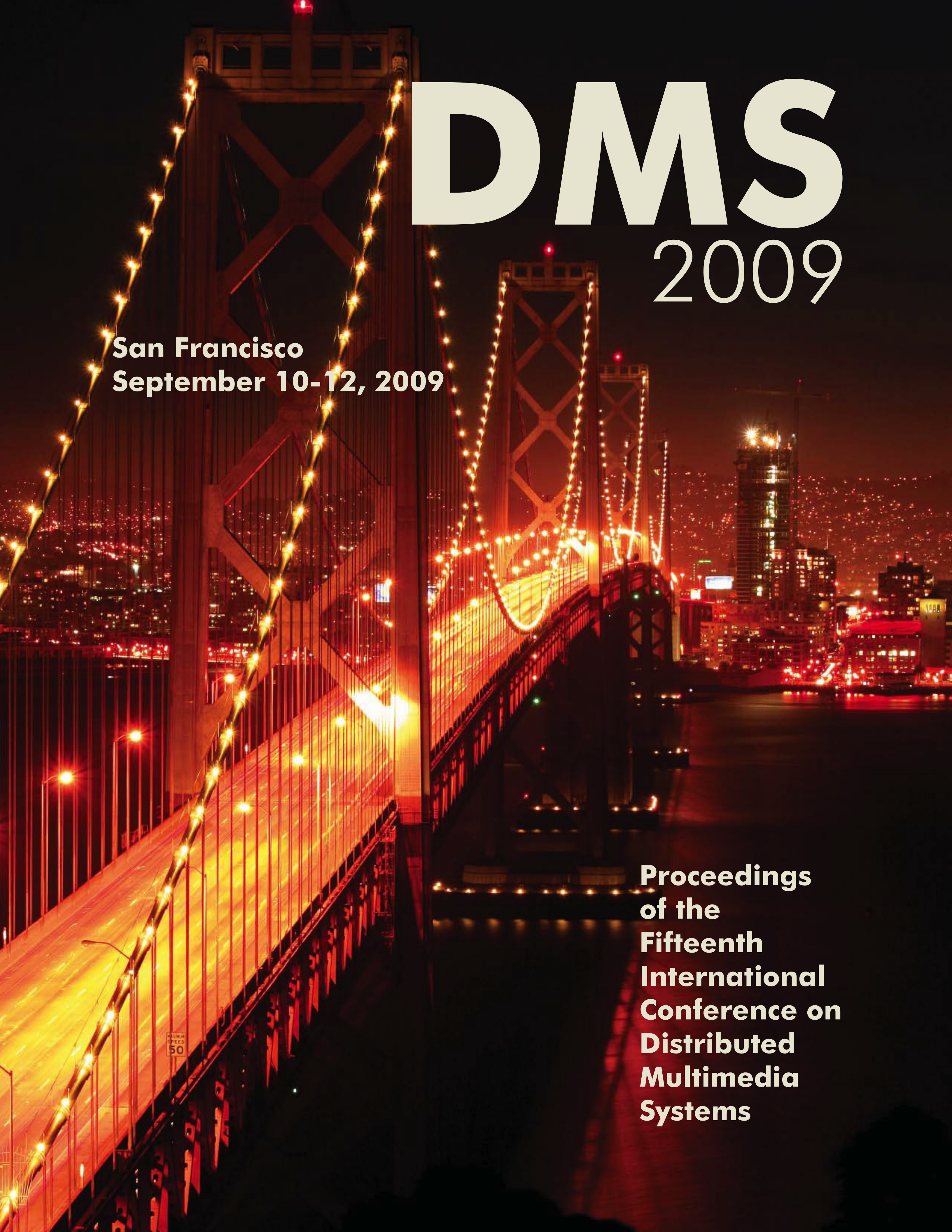


DMS 2009

San Francisco
September 10-12, 2009

Proceedings of the Fifteenth International Conference on Distributed Multimedia Systems

PROCEEDINGS

DMS 2009

The 15th International Conference on Distributed Multimedia Systems

Co-Sponsored by

Knowledge Systems Institute Graduate School, USA
Eco Controllo SpA, Italy
University of Salerno, Italy
University Ca' Foscari in Venice, Italy

Technical Program

September 10 - 12, 2009
Hotel Sofitel, Redwood City, San Francisco Bay, USA

Organized by

Knowledge Systems Institute Graduate School

ISBN 1-891706-25-X (paper)

Additional Copies can be ordered from:
Knowledge Systems Institute Graduate School
3420 Main Street
Skokie, IL 60076, USA
Tel:+1-847-679-3135
Fax:+1-847-679-3166
Email:office@ksi.edu
http://www.ksi.edu

Proceedings preparation, editing and printing are co-sponsored by
Knowledge Systems Institute Graduate School, USA
Eco Controllo SpA, Italy

Printed by Knowledge Systems Institute Graduate School

DMS 2009 Foreword

Welcome to DMS 2009, the 15th edition of the International Conference on Distributed Multimedia Systems. In past years, the DMS series of Conferences has approached the broad field of distributed multimedia systems from several complementary perspectives: theory, methodology, technology, systems and applications. The contributions of highly qualified authors from academy and industry, in the form of research papers, case studies, technical discussions and presentation of ongoing research, collected in this proceedings volume, offer a picture of current research and trends in the dynamic fields of information technology.

The main conference themes have been organized, according to a formula consolidated during past editions, into a number of thematic tracks offering to the conference attendants and to readers a convenient way to explore this vast amount of knowledge in an organized way. Two additional workshops which extended the main conference offerings, and completed the conference program (the International Workshop on Distance Education Technologies, DET 2009, and the International Workshop on Visual Languages and Computing, VLC 2009, are included here for reference.

The selection of the papers to be presented at the DMS conference this year, and to the two workshops, was based upon a rigorous review process, with an acceptance rate of about 40% of submissions received in the category of full research papers. Short papers reporting ongoing research activities and applications completed the conference content, playing the role of fostering timely discussions among the participants, not only on consolidated research achievements, but also on ongoing ideas and experiments.

Twenty-three countries are represented this year: Austria, Brazil, Canada, China, Czech Republic, France, Germany, India, Italy, Japan, Jordan, Lebanon, Malaysia, Myanmar, New Zealand, Portugal, Spain, Sweden, Switzerland, Taiwan, United Kingdom, United States, and Vietnam, giving a truly “distributed” atmosphere to the conference itself.

As program co-chairs, we appreciate having the opportunity to bring out this new edition of proceedings. We acknowledge the effort of the program committee members in reviewing the submitted papers under very strict deadlines, and the valuable advice of the conference chairs Masahito Hirakawa and Erland Jungert. Daniel Li has given excellent support by promptly replying to our requests for information about organization and technical issues. The excellent guidance of Dr. S.K. Chang has led to the success of this whole process, and we take this opportunity to thank him once again.

Finally, we thank Eco Controllo SpA, Italy, for sponsoring in part the printing of the Proceedings, the University of Salerno, Italy for sponsoring the keynote by Gennady Andrienko, and the Computer Science Department of Università Ca’ Foscari in Venice, Italy, for the financial support of one of the program co-chairs.

Augusto Celentano and Atsuo Yoshitaka
DMS 2009 Program Co-Chairs

The 15th International Conference on Distributed Multimedia Systems (DMS 2009)

September 10-12, 2009
Hotel Sofitel, Redwood City, San Francisco Bay, USA

Organizers & Committees

Steering Committee Chair

Shi-Kuo Chang, *University of Pittsburgh, USA*

Conference Co-Chairs

Masahito Hirakawa, *Shimane University, Japan*
Erland Jungert, *Linkoping University, Sweden*

Program Co-Chairs

Augusto Celentano, *Universita Ca Foscari di Venezia, Italy*
Atsuo Yoshitaka, *JAIST, Japan*

Program Committee

Vasu Alagar, *Concordia University, Canada*
Frederic Andres, *National Institute of Informatics, Japan*
Arvind K. Bansal, *Kent State University, USA*
Ioan Marius Bilasco, *Laboratoire dInformatique de Grenoble (LIG), France*
Yeim-Kuan Chang, *National Cheng Kung University, Taiwan*
Ing-Ray Chen, *Virginia Tech (VPI&SU), USA*
Shu-Ching Chen, *Florida International University, USA*
Cheng-Chung Chu, *Tunghai University, Taiwan*
Gennaro Costagliola, *Univ of Salerno, Italy*
Alfredo Cuzzocrea, *University of Calabria, Italy*

Andrea De Lucia, *Univ. of Salerno, Italy*
Alberto Del Bimbo, *Univ. of Florence, Italy*
David H. C. Du, *Univ. of Minnesota, USA*
Jean-Luc Dugelay, *Institut EURECOM, France*
Larbi Esmahi, *National Research Council of Canada, Canada*
Ming-Whei Feng, *Institute for Information Industry, Taiwan*
Daniela Fogli, *Universita degli Studi di Brescia, Italy*
Farshad Fotouhi, *Wayne State University, USA*
Alexandre Francois, *Tufts University*
Kaori Fujinami, *Tokyo University of Agriculture and Technology, Japan*
Moncef Gabbouj, *Tampere University of Technology, Finland*
Ombretta Gaggi, *Univ. of Padova, Italy*
Richard Gobel, *FH Hof, Germany*
Stefan Goebel, *ZGDV Darmstadt, Germany*
Forouzan Golshani, *Wright State University, USA*
Jivesh Govil, *Cisco Systems Inc., USA*
Angela Guercio, *Kent State University, USA*
Niklas Hallberg, *FOI, Sweden*
Hewijin Christine Jiau, *National Cheng Kung University, Taiwan*
Joemon Jose, *University of Glasgow, UK*
Wolfgang Klas, *University of Vienna, Austria*
Yau-Hwang Kuo, *National Cheng Kung University, Taiwan*
Jen Juan Li, *North Dakota State University, USA*
Fuhua Lin, *Athabasca University, Canada*
Alan Liu, *National Chung Cheng Univeristy, Taiwan*
Chien-Tsai Liu, *Taipei Medical College, Taiwan*
Chung-Fan Liu, *Kun Shan University, Taiwan*
Jonathan Liu, *University of Florida, USA*
Andrea Marcante, *Universita degli Studi di Milano, Italy*
Sergio Di Martino, *Universita degli Studi di Napoli Federico II, Italy*
Piero Mussio, *Universita degli Studi di Milano, Italy*
Paolo Nesi, *University of Florence, Italy*
Vincent Oria, *New Jersey Institute of Technology, USA*
Sethuraman Panchanathan, *Arizona State Univ., USA*
Antonio Piccinno, *Univ. of Bari, Italy*
Sofie Pilemalm, *FOI, Sweden*
Fabio Pittarello, *University of Venice, Italy*
Giuseppe Polese, *University of Salerno, Italy*
Syed M. Rahman, *Minnesota State University, USA*
Monica Sebillo, *Universita di Salerno, Italy*
Timothy K. Shih, *National Taipei University of Education, Taiwan*
Peter Stanchev, *Kettering University, USA*
Genny Tortora, *University of Salerno, Italy*
Joseph E. Urban, *Texas Tech University, USA*

Athena Vakali, *Aristotle University, Greece*
Ellen Walker, *Hiram College, USA*
KaiYu Wan, *East China Normal University, China*
Chi-Lu Yang, *Institute for Information Industry, Taiwan*
Kang Zhang, *University of Texas at Dallas, USA*

Publicity Co-Chairs

KaiYu Wan, *East China Normal University, China*
Chi-Lu Yang, *Institute for Information Industry, Taiwan*

Proceedings Cover Design

Gabriel Smith, *Knowledge Systems Institute Graduate School, USA*

Conference Secretariat

Judy Pan, *Chair, Knowledge Systems Institute Graduate School, USA*
Omasan Etuwewe, *Knowledge Systems Institute Graduate School, USA*
Dennis Chi, *Knowledge Systems Institute Graduate School, USA*
David Huang, *Knowledge Systems Institute Graduate School, USA*
Daniel Li, *Knowledge Systems Institute Graduate School, USA*

International Workshop on Distance Education Technologies (DET 2009)

September 10-12, 2009
Hotel Sofitel, Redwood City, San Francisco Bay, USA

Organizers & Committees

Workshop Co-Chairs

Tim Arndt, *Cleveland State University, USA*
Heng-Shuen Chen, *National Taiwan University, Taiwan*

Program Co-Chairs

Paolo Maresca, *University Federico II, Napoli, Italy*
Qun Jin, *Waseda University, Japan*

Program Committee

Giovanni Adorni, *University of Genova, Italy*
Tim Arndt, *Cleveland State University, USA*
Heng-Shuen Chen, *National Taiwan University, Taiwan*
Yuan-Sun Chu, *National Chung Cheng University, Taiwan*
Luigi Colazzo, *University di Trento, Italy*
Rita Francese, *University of Salerno, Italy*
Wu He, *Old Dominion University, USA*
Pedro Isaias, *Open University, Portugal*
Qun Jin, *Waseda University, Japan*
Paolo Maresca, *University Federico II, Napoli, Italy*
Syed M. Rahman, *Minnesota State University, USA*
Teresa Roselli, *University of Bari, Italy*
Nicoletta Sala, *University of Italian Switzerland, Switzerland*
Giuseppe Scanniello, *University of Salerno, Italy*
Hui-Kai Su, *Nanhua University, Taiwan*

Yu-Huei Su, *National HsinChu University of Education, Taiwan*
Kazuo Yana, *Hosei University, Japan*

International Workshop on Visual Languages and Computing (VLC 2009)

September 10-12, 2009
Hotel Sofitel, Redwood City, San Francisco Bay, USA

Organizers & Committees

Workshop Co-Chairs

Giuseppe Polese, *University of Salerno, Italy*
Giuliana Vitiello, *University of Salerno, Italy*

Program Chair

Gem Stapleton, *University of Brighton, UK*

Program Committee

Dorothea Blostein, *Queen's University, Canada*
Paolo Buono, *University of Bari, Italy*
Alfonso F. Cardenas, *University of California, USA*
Kendra Cooper, *University of Texas at Dallas, USA*
Maria Francesca Costabile, *University of Bari, Italy*
Gennaro Costagliola, *University of Salerno, Italy*
Philip Cox, *Dalhousie University, Canada*
Vincenzo Deufemia, *University of Salerno, Italy*
Stephan Diehl, *University of Trier, Germany*
Jing Dong, *The University of Texas at Dallas, USA*
Filomena Ferrucci, *University of Salerno, Italy*
Andrew Fish, *University of Brighton, UK*
Paul Fishwick, *University of Florida, USA*
Manuel J. Fonseca, *INESC-ID, Portugal*
Dorian Gorgan, *Technical University of Cluj-Napoca, Romania*
Corin Gurr, *University of Reading, UK*
Tracy Hammond, *Texas A&M University, USA*

Maolin Huang, *University of Technology, Sydney, Australia*
Erland Jungert, *Linkoping University, Sweden*
Lars Knipping, *Technische Universitat Berlin, Germany*
Hideki Koike, *University of Electro-Communications Tokyo, Japan*
Jun Kong, *North Dokota State University, USA*
Zenon Kulpa, *Institute of Fundamental Technological Research, Poland*
Robert Laurini, *University of Lyon, France*
Benjamin Lok, *University of Florida, USA*
Kim Marriott, *Monash University, Australia*
Rym Mili, *University of Texas at Dallas, USA*
Piero Mussio, *University of Milan, Italy*
Luca Paolino, *University of Salerno, Italy*
Joseph J. Pfeiffer, *New Mexico State University, USA*
Beryl Plimmer, *University of Auckland, New Zealand*
Giuseppe Polese, *University of Salerno, Italy*
Steven P. Reiss, *Brown University, USA*
Gem Stapleton, *University of Brighton, UK*
David Stotts, *University of North Carolina, USA*
Nik Swoboda, *Universidad Politecnica de Madrid, Spain*
Athanasios Vasilakos, *University of Western Macedonia, Greece*
Giuliana Vitiello, *University of Salerno, Italy*
Kang Zhang, *University of Texas at Dallas, USA*

Table of Contents

Distributed Multimedia Computing & Networks and Systems

Digital Home and HealthCare - I

Digital Home and HealthCare - II

Distributed Multimedia Systems - II

Technologies for Digital Television

Emergency Management and Security

Distributed Multimedia Systems - III

Distributed Multimedia Systems - IV

Distributed Multimedia Systems - V

DET Workshop

Eclipse and Jazz Technologies for E-learning

Learning Objects: Methodologies, Technologies and Experiences

E-learning and The Arts

E-learning

VLC Workshop

Visual Analytics - I

Visual Languages and Environments for Software Engineering

Visual Semantics, Tools and Layout

Sketch Computing

Visual Analytics - II

Note: (S) means short paper.

Keynote I: Slow Intelligence Systems

Shi-Kuo Chang

Abstract

In this talk I will introduce the concept of slow intelligence. Not all intelligent systems are fast. There are a surprisingly large number of intelligent systems, quasi-intelligent systems and semi-intelligent systems that are slow. Such slow intelligence systems are often neglected in mainstream research on intelligent systems, but they are really worthy of our attention and emulation. I will discuss the general characteristics of slow intelligence systems and then concentrate on evolutionary query processing for distributed multimedia systems as an example of artificial slow intelligence systems.

About Shi-Kuo Chang

Dr. Chang received the B.S.E.E. degree from National Taiwan University in 1965. He received the M.S. and Ph.D. degrees from the University of California, Berkeley, in 1967 and 1969, respectively. He was a research scientist at IBM Watson Research Center from 1969 to 1975. From 1975 to 1982 he was Associate Professor and then Professor at the Department of Information Engineering, University of Illinois at Chicago. From 1982 to 1986 he was Professor and Chairman of the Department of Electrical and Computer Engineering, Illinois Institute of Technology. From 1986 to 1991 he was Professor and Chairman of the Department of Computer Science, University of Pittsburgh. He is currently Professor and Director of the Center for Parallel, Distributed and Intelligent Systems, University of Pittsburgh. Dr. Chang is a Fellow of IEEE. He published over 230 papers and 16 scientific books. He is the founder and co-editor-in-chief of the international journal, Visual Languages and Computing, published by Academic Press, the editor-in-chief of the international journal, Software Engineering & Knowledge Engineering, published by World Scientific Press, and the co-editor-in-chief of the international journal on Distance Education Technologies. Dr. Chang pioneered the development of Chinese language computers, and was the first to develop a picture grammar for Chinese ideographs, and invented the phonetic phrase Chinese input method.

Dr. Chang's literary activities include the writing of over thirty novels, collections of short stories and essays. He is widely regarded as an acclaimed novelist in Taiwan. His novel, The Chess King, was translated into English and German, made into a stage musical, then a TV mini-series and a movie. It was adopted as textbook for foreign students studying Chinese at the Stanford Center (Inter-University Program for Chinese Language Studies administered by Stanford University), Taipei, Taiwan. In 1992, Chess King was adopted as supplementary reading for high school students in Hong Kong. The short story, "Banana Boat", was included in a textbook for advanced study of Chinese edited by Neal Robbins and published by Yale University Press. University of Illinois adopted "The Amateur Cameraman" in course materials for studying Chinese. Dr. Chang is also regarded as the father of science fiction in Taiwan. Some of Dr. Chang's SciFi short stories have been translated into English, such as "City of the Bronze Statue", "Love Bridge”, and "Returning”. His SciFi novel, The City Trilogy, was published by Columbia University Press in May 2003.

Keynote II: Geographic Visualization of Movement Patterns

Gennady Andrienko and Natalia Andrienko

Abstract

We present our recent results in visualization and visual analytics of movement data. The GeoPKDD project (Geographic Privacy-aware Knowledge Discovery and Delivery) and the recently started DFG project ViAMoD (Visual Spatiotemporal Pattern Analysis of Movement and Event Data) have brought into existence an array of new methods enabling the analysis of really large collections of movement data. Some of the methods are applicable even to data not fitting in the computer main memory. These include the techniques for database aggregation, cluster-based classification, and incremental summarization of trajectories. The remaining methods can deal with data that fit in the main memory but are too big for the traditional visualization and interaction techniques. Among these methods are interactive visual cluster analysis of trajectories and dynamic aggregation of movement data. The visual analytics methods are based on the interplay of computational algorithms and interactive visual interfaces, which support the involvement of human capabilities for pattern recognition, association, interpretation, and reasoning. The projects have also moved forward the theoretical basis for visual analytics methods for movement data. We discuss analysis tasks and problems requiring further research.

About Gennady Andrienko

Gennady Andrienko received his Master degrees in Computer Science from Kiev State University in 1986 and Ph.D. equivalent in Computer Science from Moscow State University in 1992. He undertook research on knowledge-based systems at the Mathematics Institute of Moldavian Academy of Sciences (Kishinev, Moldova), then at the Institute on Mathematical Problems of Biology of Russian Academy of Science (Pushchino Research Center, Russia). Since 1997 Dr. Andrienko has a research position at GMD, now Fraunhofer Institute for intelligent Analysis- and Information Systems (IAIS). He is a co-author of the monograph "Exploratory Analysis of Spatial and Temporal Data", 30+ peer-reviewed journal papers, 10+ book chapters, and 100+ papers in conference proceedings. He has been involved in numerous international research projects. His research interests include geovisualization, information visualization with a focus on spatial and temporal data, visual analytics, interactive knowledge discovery and data mining, spatial decision support and optimization.

Proceedings
International Conference on Distributed Multimedia Systems
(DMS 2009)

Co-Editors

Augusto Celentano, *Universita Ca Foscari di Venezia, Italy*
Atsuo Yoshitaka, *JAIST, Japan*

Demonstrating the Effectiveness of Sound Spatialization in Music and Therapeutic Applications

Masahito Hirakawa, Mirai Oka, Takayuki Koyama[1], and Tetsuya Hirotomi
Interdisciplinary Faculty of Science and Engineering, Shimane University, Japan
{hirakawa, hirotomi}@cis.shimane-u.ac.jp

Abstract

Most of the existing computer systems express information visually. While vision plays an important role in interaction between the human and the computer, it is not the only channel. We have been investigating a multimedia system which is capable of controlling the spatial position of sounds on a two-dimensional table.

In this paper we present applications of the system to sound mashup and reminiscence/life review, in order to demonstrate the effectiveness of sound spatialization in collaborative work environments. Users can collaborate with each other with the help of sound objects which are spatialized on the table, in addition to graphical images.

1. Introduction

Multimedia is a basis of modern computers. In fact a variety of studies have been investigated so far. Graphical user interfaces, or visual languages in a broader sense, are one such example toward development of advanced computers in the early days of multimedia research. Since humans are sensitive to vision, it is natural that our attention had been paid to the use of visual information in interaction between the user and the computer.

Meanwhile, audition is another important channel for interaction. The idea of so-called earcon [1] was first proposed to present specific items or events by means of abstract patterns in loudness, pitch, or timbre of sounds. Studies of auditory interface have been done actively in such applications as menu navigation [2], mobile service notifications [3], [4], mobile games [5], and human movement sonification [6].

In those trials, sound patterns or notes are a matter of concern. While they give the user a great impact in understanding the associated events, the spatial position of sounds influences the user's understanding as well [7].

Stereo and 5.1-channel surround systems which have been used widely make it possible for the listener to feel the sound position. It should be mentioned that, however, the best spot for listening is fixed in those settings. If the listener is out of the spot, a reality of the sound space cannot be maintained any more. Due to this fact, those systems are suitable for the application where a limited number of listeners sit in a limited space.

In collaborative or multi-user computing environments, the system should support a mechanism that each of the users can catch where sounds are placed, irrelevant to his/her standing position and direction.

The authors have investigated a tabular sound system for a couple of years [8], [9]. The system is equipped with a meter square table in which 16 speakers are placed in a 4 x 4 grid layout. Multiple sound streams can be presented simultaneously by properly controlling the loudness for those speakers. Additionally, computer generated graphical images are projected on its surface. We call this table "Sound Table." Users who surround the table can feel spatial sounds with the associated images. In addition, a special stick-type input device is provided for specification of commands. It is important to note that the users do not need to wear any special devices for interacting with the system.

In this paper we present applications of the system to sound mashup and reminiscence/life review, in order to demonstrate the effectiveness of sound spatialization in collaborative work environments.

[1] Mr. Koyama is now with ICR, Japan.

2. Related Work

Sound spatialization studies have been active in a human-computer interaction domain [10]. One practical example is a computer game named “Otogei” which was produced by Bandai. The player wears a headphone and tries to attack the approaching enemies by relying on a stereo sound. [11] - [13] presented sound-based guidance systems which guide a user to a desired target location by varying the loudness and balance of a sound played. There exist some other approaches of using sounds for assistance of, for example, car driving [14], mail browsing [15], geographical map navigation [16], and object finding in 3D environments [17].

Here, those systems assume a headphone or a specially designed hardware as an interaction device. A user is separated from others, and each of the users hears a different sound even though multiple people participate in a common session. This feature is advantageous in some cases, but not recommended for collaborative work environments.

[18] conducted experiments on the use of non-speech audio at an interactive multi-user tabletop display under two different setups. One is a localized sound where each user has his or her own speaker, and the other is a coded sound where users share one speaker but waveforms of the sounds are varied so that a different sound is played for each user. This approach could be one practical solution to business-oriented applications, but is not sufficient for sound-centric applications (e.g. computer music).

Transition Soundings [19] and Orbophone [20] are specialized interfaces using multiple speakers for interactive music making. A large number of speakers are mounted in a wall-shaped board in Transition Soundings, while Orbophone houses multiple speakers in a dodecahedral enclosure. Both systems are deployed for sound art.

Other related approaches of using multi-channel speakers appear in [21], [22]. While they provide sophisticated functionality, their system setting is rather complex and specialized. As will be explained in the next section, we use conventional speakers and sound boards and no specialized hardware is used at all.

3. Tabular Sound Spatialization System

The sound spatialization system [8], [9] we have developed as a platform for sound-based collaborative applications is organized by Sound Table as its central equipment, a pair of cameras, a video projector, and a PC. Figure 1 shows its physical setup (The PC is not shown).

Figure 1. Sound spatialization system

Sound Table is a physical table in which 16 speakers are equipped in a 4 x 4 matrix, as shown in Fig. 2. It is of 90cm width and depth, and 73cm height. Two 8-channel audio interfaces (M-AUDIO FireWire 410) are equipped to the PC, and connected to Sound Table through a 16-channel amplifier. Multiple sounds can be output at one time at different positions.

Figure 2. Sound Table

We have analyzed how accurate the sound positioning is through experiments, that is, errors in distance between the simulated sound position and the perceived sound position. The average error of sound position identification for moving sounds is 0.52 in horizontal direction and 0.72 in depth direction, where their values are normalized by the distance between two adjacent speakers (24cm). Further details are given in [8].

The surface of Sound Table is covered by a white cloth so that computer-generated graphical images are projected onto it. Multiple users can interact with the system through both auditory and visual channels.

A stick-type input device whose base unit is Nintendo Wii Remote is provided as shown in Fig. 3. Position and posture of the device in a 3D space over Sound Table are captured by the system, as well as button press.

Figure 3. Sticky input device

Meanwhile, a task of identifying user's gestures which include tap, sting and release, attack, flick, and tilt is separated from that of interpreting their semantic meanings in a certain application so that application development can be made easier. We adopted the OSC (Open Sound Control) protocol for communication of messages among processing modules. For details of the software development framework, please refer to [9].

We have first implemented a simple music application of the system in order to demonstrate its functionality [9]. Here, in this paper, we will show more practical applications at which the sound spatialization facility plays a significant role.

4. Music Mashup Application

4.1 Background

In computer music, people are interested in creating and performing music. Musical instruments which are augmented by digital technologies have been proposed. TENORI-ON [23] is an example. It gives a "visible music" interface at which a 16x16 matrix of touch sensible LED switches allows a user to play music intuitively. Some researchers put emphasis on the instrument part (e.g., [24], [25]), but some others focused attention on user-interface where tactile, gestural, or multimodal features are emphasized (e.g., [26], [27]).

The trials mentioned above focus on interactive music composition. There have been few trials allowing the user to enjoy manipulating the spatial position of sound sources (e.g., virtual music performers), while it is of a great importance to people in order to attain reality [10]. Pinocchio [28] and the one exhibited at Sony ExploraScience museum are examples which emphasize localization of sound.

Meanwhile, the online music software Massh! [29] inspired us with its distinguished functionality and interactive features. It enables users mix sound samples or loops to make a new song (i.e., mashup). Furthermore, its visual user interface is highly interactive. Sound loops are graphically represented on the screen as rotating circular waveforms. They can form a group (i.e. mix), which are played in sync with each other.

Sound loops are presented in Massh! as visual clues, but no sound spatialization is available. We consider adding a sound spatialization facility for more attractive music mashup.

4.2 Design policy

Several different interface designs for music mashup in our system setting can be thought.

One possibility is that, considering music loops are time-based media and their execution (play) is limited to one part within the whole at a time, a music loop/sample is represented in a form of timeline with a slider showing which part of the music loop/sample is being played. Multiple sliders may be assigned to one music loop/sample, allowing the player to have a composition that employs a melody with one or more imitations of the melody played after a given duration, that is, a canon.

Meanwhile, we take another approach where music is organized by multiple moving sound objects which correspond to sound samples or loops. While no play position control is available for the objects, flexibility is given to them in respect of their moving paths. This fits well to our system architecture.

Here, in order to have variation of sounds generated, we prepare two path patterns: a straight line and a circular line. Multiple sound objects may be associated with one path. When a sound object comes to a crossing point where two or more paths are overlapped, the object may change its path to another.

4.3 Implementation

We have built an actual music mashup application on the tabular sound spatialization system.

First, the user determines a path for sound object(s) on the table by placing certain gestures as explained below.

For specification of a straight line, the user touches the stick device at a starting point on the table, and then brings it to a desirable terminal position with keeping its head on the table. The straight line has a handle in a triangular shape at each side of the line (see Fig. 4). The user can change the length and angle of a line by manipulating its handle.

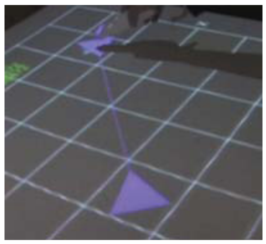
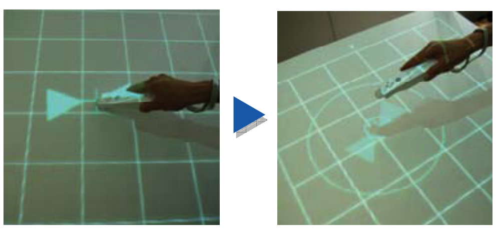


Figure 4. Specification of a straight line path

On the other hand, a circular line can be generated by bringing one handle of a predefined straight line close to the other handle, as shown in Fig. 5(a). The user is allowed to modify the position and size of a circular line by dragging a center marker and a special marker on the line, respectively (see Fig. 5(b)).



(a) (b)

Figure 5. Specification of a circular line path

Music starts by generating sound objects on the table. Generation of a sound object is carried out by tilting the stick device while pushing a button of the device. Graphically, a sound object takes a circular shape with a certain color and size. The color corresponds to a sound sample/loop, while the size corresponds to its loudness. The size of a sound object is determined, when the object is instanciated, by the position of the stick device in a 3D space on the table. Higher the spatial position, larger the circle size and thus louder the generated sound.

When a sound object is placed on a line, it starts moving along the line. Users enjoy feeling the movement of the sound. In the present implementation, change of the path from one line to another at a crossing point happens by a certain possibility. Figure 6 shows such examples. Furthermore, it is allowed for the user to take a sound object to another position after its creation. If the object is placed at the position where no path line exists, it keeps its position and doesn’t move.

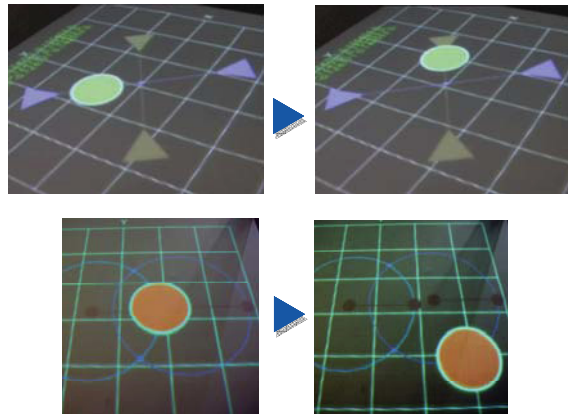


Figure 6. Change of a path

Meanwhile, when the device is swung down over a sound object which is sounding, the sound is terminated. At the same time, its color becomes black to see the change. If the gesture is applied again, the object restarts sounding.

When the user places a gesture of handling a sound object on the table with a light quick blow, it flows out of the table with graphical effects - i.e., the object is broken in segments. Semantically this means deletion of the object.

Figure 7 shows a snapshot of the system in use. Multiple users can play collaboratively each other. Manipulation of sound objects and lines which may be specified by other users brings a change of sounds in real time. This notifies each user of others’ play, and stimulates him/her to have a reaction.

Figure 7. Collaborative play with the system

Having a feeling of sound movement is attractive and fun in such a music application realized in our trial. Here we noticed the importance of authoring effective content to give users better impression in their performance. Experimental evaluation of the usefulness of the proposed music application still remains.

5. Supporting Reminiscence for Older People

5.1 Background

Reminiscence therapy is a psychosociological therapeutic approach to the care of older people [30], [31]. Older people recall various experiences from their past life and share them with others to facilitate pleasure, quality of life, emotional stability or adaptation to present circumstances, and to reduce isolation and depression.

In practice, due to a rapid increase in the elderly population, interest in reminiscence therapy has continued to grow. Trials have actually been carried out in hospitals, day care, nursing homes, and other settings, where reminiscence therapy is usually conducted in a group guided by an experienced staff.

Meanwhile, in a reminiscence session, the staff shows visual media such as photographs and pictures as a clue. Some other media including music, smell, and tactile may be used as well to make the session successful. [32] and [33] present computer-based multimedia conversion aids in which audio, video, animation and/or QuickTime VR are utilized.

It is noted that, in the existing trials of reminiscence therapy with music, songs or melodies are a matter of concern. It is expected that the position of sounds and its movement work considerably to help people in recalling experiences and then initiating their speech.

5.2 Design policy

We consider that there are two key points in the development of a computer-assisted system implementing reminiscence for practical use.

One is the friendliness and effectiveness of the system to participants (older people). They are not willing to use a computer and, thus, its user-interface should be natural and simple.

The other concerns the utility of the system for an experienced staff who guides older people in reminiscence. There are demands of helping him/her in creation of reminiscence materials and gathering of data which are useful for analysis of the session, for example, how long each of the participants spoke and which topics he/she was interested in.

In this trial, we consider issues of a multimodal interface for creation and play of reminiscence materials. A facility of recoding and analyzing activities presented by participants will be reported elsewhere.

The interface needs providing a facility to place a sound at any position on a picture and specify its arbitrary movement on the picture as, for example, a child runs around in a playground. The specification should be understandable so that the user can edit it. Here, simplicity is a matter of vital importance in its design.

5.3 Implementation

For creation of a reminiscence material, the staff first selects a picture from a database, and then assigns sound objects on it by manipulating the sticky device. Each of the created sound objects is visualized in an icon so that the staff can easily identify the position and some other states of the object, as shown in Fig. 8. Those states include the mobility (moving object or stable object) and sound existence (on or off). When the staff drags the icon (i.e., sound object) by using the sticky device on the table, its path is recorded as traversed. He/she may repeat the tasks explained above to define a complete set of the reminiscence material.

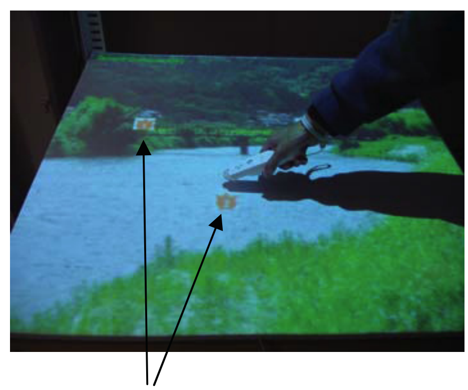


Figure 8. Assignment of sound objects

Once the specification is completed, it is ready to play the material. The staff can switch from one picture (with sounds) to another by pressing a button of the sticky device. Icons as sound markers are not displayed anymore during the playback.

Meanwhile, a preliminary evaluation of the system has been conducted. A group of three university students participated in the test where they were asked to have a reminiscence session using the system. We compared system performances in two settings (with and without sounds) by a questionnaire with three questions: "Was the communication lively?", "Was it helpful to initiate a speech?", and "Which setting is advantageous?".

All of the subjects marked higher score to the setting with sounds than that with no sounds. In addition, the following opinions are given by the subjects.

- The session with sounds stimulated reminiscence.
- Combination of background sounds with foreground sounds, which are listened to consciously, would be beneficial.

Though further detailed experiments must be conducted, this system setup would be of help in performing reminiscence therapy. Usefulness of the authoring facilities for the experienced staff needs to be investigated.

By the way, in the current implementation, we assume static images. We will investigate an extension so that videos can be used as a medium for reminiscence therapy. The system should then provide a facility that sound objects follow target objects in a video. Of course an experienced staff doesn't want to learn complex operations in authoring. It is necessary to design an interface so as not to make the authoring of such dynamic content difficult. A mechanism of video editing based on object movement that one of the authors proposed before [34] would be helpful to the development.

6. Conclusions

We investigated in the paper how actually sound positioning serves us as an effective technique for implementation of advanced computer applications. As practical examples, two applications to music mashup and reminiscence were presented, which have been implemented on top of the tabular sound spatialization system we developed before. Users can collaborate with each other with the help of sound objects which are spatialized on the table, in addition to graphical images.

Further studies still remain, which include synchronization of sound objects running on a certain path as to music mashup, and user tests by older people in reminiscence.

Acknowledgement

This work has been supported in parts by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research, 20500481, 2008.

References

[1] M. M. Blattner, "Multimedia Interface Design", Addison-Wesley Pub., 1992.

[2] P. Yalla and B. N. Walker, "Advanced Auditory Menus: Design and Evaluation of Auditory Scroll Bars," Proc., Int'l ACM SIGACCESS Conf. on Computers and Accessibility, pp.105-112, 2008.

[3] S. Garzonis, C. Bevan, and E. O'Neill, "Mobile Service Audio Notifications: Intuitive Semantics and Noises," Proc., ACM Australasian Conf. on Computer-Human Interaction: Designing for Habitus and Habitat, pp.156-163, 2008.

[4] E. Hoggan and S. Brewster, "Designing Audio and Tactile Crossmodal Icons for Mobile Devices," Proc., ACM Int'l Conf. on Multimodal Interfaces, pp.162-169, 2007.

[5] I. Ekman, L. Ermi, J. Lahti, J. Nummela, P. Lankoski, and F. Mäyrä, "Designing Sound for a Pervasive Mobile Game," Proc., ACM SIGCHI Int'l Conf. on Advances in Computer Entertainment Technology, pp.110-116, 2005.

[6] A. O. Effenberg, "Movement Sonification: Effects on Perception and Action," IEEE MultiMedia, Vol.12, No.2, pp.53-59, Apr.-June 2005.

[7] J. J. Nixdorf and D. Gerhard, "RITZ: A RealTime Interactive Tool for Spatialization", Proc., ACM Int'l Conf. on Multimedia, pp.687-690, 2006.

[8] T. Nakaie, T. Koyama, and M. Hirakawa, "Development of a Collaborative Multimodal System with a Shared Sound Display", Proc., IEEE Conf. on Ubi-Media Computing, pp.14-19, 2008.

[9] T. Nakaie, T. Koyama, and M. Hirakawa, "A Table-based Lively Interface for Collaborative Music Performance", Proc. Int'l Conf. on Distributed Multimedia Systems, pp.184-189, 2008.

[10] H. J. Song and K. Beilharz, "Aesthetic and Auditory Enhancements for Multi-stream Information Sonification", Proc., Int'l Conf. on Digital Interactive Media in Entertainment and Arts, pp.224-231, 2008

[11] S. Strachan, P. Eslambolchilar, R. Murray-Smith, "gpsTunes Controlling Navigation via Audio Feedback", Proc., ACM MobileHCI'05, pp.275-278, 2005.

[12] M. Jones and S. Jones, "The Music is the Message", ACM interactions, Vol.13, No.4, pp.24-27, July&Aug. 2006.

[13] J. Dodiya and V. N. Alexandrov, "Use of Auditory Cues for Wayfinding Assistance in Virtual Environment: Music Aids Route Decision," Proc., ACM Symp. on Virtual Reality Software and Technology, pp.171-174, 2008.

[14] J..Sodnik, S..Tomazic, C..Dicke, and M. Billinghurst, “Spatial Auditory Interface for an Embedded Communication Device in a Car,” Proc., IEEE Int’l Conf. on Advances in Computer-Human Interaction, pp.69-76, 2008.
[15] D. I. Rigas and D. Memery, “Multimedia E-Mail Data Browsing: The Synergistic Use of Various Forms of Auditory Stimuli,” Proc., IEEE Int’l Conf. on Information Technology: Computers and Communications, pp.582-588, 2003.
[16] H. Zhao, B. K. Smith, K. Norman, C. Plaisant, and B. Shneiderman, “Interactive Sonification of Choropleth Maps,” IEEE MultiMedia, Vol.12, No.2, pp.26-35, Apr.-June 2005.
[17] K. Crommentuijn and F. Winberg, “Designing Auditory Displays to Facilitate Object Localization in Virtual Haptic 3D Environments,” Proc., Int’l ACM SIGACCESS Conf. on Computers and Accessibility, pp.255-256, 2006.
[18] M. S. Hancock, C. Shen, C. Forlines, and K. Ryall, “Exploring Non-Speech Auditory Feedback at an Interactive Multi-User Tabletop”, Proc., Graphics Interface 2005, pp.41-50, 2005.
[19] D. Birchfield, K. Phillips, A. Kidane, and D. Lorig, “Interactive Public Sound Art: A Case Study”, Proc., Int’l Conf. on New Interfaces for Musical Expression, pp.43-48, 2006.
[20] D. Lock and G. Schiemer, “Orbophone: A New Interface for Radiating Sound and Image”, Proc., Int’l Conf. on New Interfaces for Musical Expression, pp.89-92, 2006.
[21] T. Ogi, T. Kayahara, M. Kato, H. Asayama, and M. Hirose, “Immersive Sound Field Simulation in Multi-screen Projection Displays”, Proc., Eurographics Workshop on Virtual Environments, pp.135-142, 2003.
[22] C. Ramakrishnan, J. Goßmann, and L. Brümmer, “The ZKM Klangdom”, Proc., Int’l Conf. on New Interfaces for Musical Expression, pp.140-143, 2006.
[23] http://www.global.yamaha.com/design/tenori-on/
[24] Y. Takegawa, M. Tsukamoto, T. Terada, S. Nishio, “Mobile Clavier: New Music Keyboard for Flexible Key Transpose”, Proc., Int’l Conf. on New Interfaces for Musical Expression, pp.82-87, 2007.
[25] D. Overholt, “The Overtone Violin”, Proc. Int’l Conf. on New Interfaces for Musical Expression, pp.34-37, 2005.
[26] S. Jorda, G. Geiger, M. Alonso, and M. Kaltenbrunner, “The reacTable: Exploring the Synergy between Live Music Performance and Tabletop Tangible Interfaces”, Proc., ACM Conf. on Expressive Character of Interaction, pp.139-146, 2007.
[27] A. Crevoisier, C. Bornand, S. Matsumura, and C. Arakawa, “Sound Rose: Creating Music and Images with a Touch Table”, Proc., Int’l Conf. on New Interfaces for Musical Expression, pp.212-215, 2006.
[28] B. Bruegg, C. Teschner, P. Lachenmaier, E. Fenzl, D. Schmidt, and S. Bierbaum, “Pinocchio: Conducting a Virtual Symphony Orchestra”, Proc., ACM Int’l Conf. on Advances in Computer Entertainment Technology, pp.294-295, 2007.
[29] N. Tokui, “Massh! - A Web-based Collective Music Mashup System”, Proc., Int’l Conf. on Digital Interactive Media in Entertainment and Arts, pp.526-527, 2008.
[30] R. N. Butler, “Age, Death, and Life Review”, Living With Grief: Loss in Later Life (Ed. by K. J. Doka), Hospice Foundation of America, 2002.
[31] Y. C. Lin, Y. T. Dai, and S. L. Hwang, “The Effect of Reminiscence on the Elderly Population: A Systematic Review”, Public Health Nursing, Vol.20, No.4, pp.297-306, Aug. 2003.
[32] N. Alm, R. Dye, G. Gowans, J. Campbell, A. Astell, and M. Ellis, “A Communication Support System for Older People with Dementia”, IEEE Computer, Vol.40, No.5, pp.35-41, May 2007.
[33] N. Kuwahara, S. Abe, K. Yasuda, and K. Kuwabara , “Networked Reminiscence Therapy for Individuals with Dementia by Using Photo and Video Sharing”, Proc., Int’l ACM Conf. on Computers and Accessibility, pp.125-132, 2006.
[34] Y. Wang and M. Hirakawa, "Video Editing Based on Object Movement and Camera Motion", Proc., ACM Int’l Working Conf. on Advanced Visual Interfaces, pp.108-111, 2006.

End-User Development in the Medical Domain

Maria Francesca Costabile*, Piero Mussio°, Antonio Piccinno*,
Carmelo Ardito*, Barbara Rita Barricelli°, Rosa Lanzilotti*
**Dipartimento di Informatica, Università di Bari, ITALY*
°DICO, Università di Milano, ITALY
{costabile, piccinno, ardito, lanzilotti}@di.uniba.it, {mussio, barricelli}@dico.unimi.it

Abstract

Nowadays, users are evolving from consumers of content and tools to producers of them, also becoming co-designers of their tools and content. In this paper we report on a methodology that supports this evolution. It derives from our experience in participatory design projects to develop multimedia systems to be used by professional people in their work practice, supporting these people not only in performing activities in their specific domain, but also allowing them to tailor their virtual tools and environments and even to create and modify software artifacts. The latter are defined activities of End-User Development (EUD). We show in this paper why EUD is particularly needed in the medical domain and how the methodology we have defined can be successfully applied to this domain.

1. Introduction

A significant evolution of HCI practice is now underway. Users are evolving from consumers of content and tools to producers of them, increasingly becoming co-designers of their tools and content [1, 2]. This evolution poses problems to software designers, because users require software environments to create their own tools empowered by the software but not being obliged to become software experts. New methodologies arise which support this evolution.

In this paper, we report on a methodology rising from our experience in participatory design projects to develop multimedia systems to support professional people in their work practice. We illustrate our approach by considering distributed multimedia systems in the medical domain. Besides physicians, in the last years we cooperated with other communities of professional people, such as geologists and mechanical engineers. These communities have some common characteristics and requirements: a) they all perform their activities as competent practitioners, in that "they exhibit a kind of knowing in practice, most of which is tacit" and they "reveal a capacity for reflection on their intuitive knowing in the midst of action and sometimes use this capacity to cope with the unique, uncertain, and conflicted situations of practice" [3]; b) they are experts in a specific discipline (e.g. medicine, geology, etc.), not necessarily experts in computer science. They use their wisdom and knowledge in performing their activities, need to collect and share the knowledge they create to achieve their goals. Thus, they are knowledge workers who need to become producers of content and software tools.

The research we carried out in the last few years is devoted to design and development of multimedia interactive systems that support people in performing activities in their specific domains, but also allow them to tailor these environments so that they can better adapt to their needs, and even to create or modify software artefacts. The latter are defined activities of End-User Development (EUD) [1, 2]. By end users we mean people who use computer systems as part of daily life or daily work, but are not interested in computers per se [1, 4].

We show in this paper why End-User Development (EUD) is particularly needed in the medical domain and how the methodology we have defined to support EUD can be successfully applied to this domain.

2. The overall approach

In the years, we have been developing an approach to participative design and to the creation of software infrastructures that support EUD activities as well as knowledge creation and sharing performed by knowledge workers in a specific domain.

The approach capitalizes on the model of the HCI process and on the theory of visual sentences we have developed [5]. HCI is modeled as a *syndetic, holistic*, *dynamic* process: syndetic in that it is a process in which two systems of different nature (the cognitive human and the computational machine) cooperate in the development of activities; holistic in that it is a process whose behavior emerges from the behaviors of the two systems, and cannot be foreseen in advance; dynamic in that the HCI process occurs through the cyclical exchange of messages (e.g. visual, audio or haptic messages) between human and machine in a temporal sequence. Each message exchanged between the two communicants is subject to two interpretations: one performed by the human and one performed by the computer, based on the code created by the program designer [1].

The research resulted in the definition of the *Software Shaping Workshop* (SSW) methodology [1], which adopts a participatory approach that allows a team of experts, including at least software engineers, HCI experts and end users to cooperate in the design and implementation of interactive systems. The aim of this methodology is to create systems that are easily understood by their users because they "speak" users' languages. Such systems are based on an infrastructure constituted by software environments, called Software Shaping Workshops (SSW or briefly workshops), and communication channels among these workshops. The term *workshop* comes from the analogy with an artisan or engineer workshop, i.e. the workroom where a person finds all and only those tools necessary to carry out her/his activities. Following the analogy, SSWs are virtual workshops in which users shape their software tools. Each adopts a domain-oriented interaction language tailored to its user's culture, in that it is defined by evolving the traditional user notations and system of signs.

End users, as knowledge workers, interact with SSWs to perform their activities, to create and share knowledge in their specific domains, to participate in the design of the whole system, even at use time. Indeed, End-User Development (EUD) implies the active participation of end users in the software development process allowing users to create and/or modify software artefacts. In this perspective, tasks that are traditionally performed by professional software developers are transferred to end users, who need to be specifically supported in performing these tasks. Some EUD-oriented techniques have already been adopted by software for the mass market, such as the adaptive menus in MS Word™ or some "Programming by Example" techniques in MS Excel™. However, we are still quite far from their systematic adoption.

To permit EUD activities, we defined a meta-design approach that distinguishes two phases: the first phase consisting in designing the design environment (meta-design phase), the second one consisting in designing the actual applications by using the design environment. The two phases are not clearly distinct and are executed several times in an interleaved way, because the design environments evolve both as a consequence of the progressive insights the different stakeholders gain into the design process and as a consequence of the feedbacks provided by end users working with the system in the field [1, 2].

The methodology offers to each expert (software engineers, HCI experts, end users as domain experts) a software environment (SSW), by which the expert contributes to shape software artefacts. In this way the various experts, each one through her/his SSW, can access and modify the system of interest according to her/his own culture, experience, needs, skills. They can also exchange the results of these activities to converge to a common design. The proposed approach fosters the collaboration among communities of end users, managers, and designers, with the aim of increasing motivation and reducing cognitive and organizational cost, thus providing a significant contribution to EUD's evolution.

The SSW infrastructure resulting from the application of the SSW methodology is a network of interactive environments (software workshops) which communicate through the exchange of annotations and boundary objects. In particular, the prototype of the application being developed is used as a boundary object, which can be used and annotated by each stakeholder [6]. Each stakeholder participates to the design, development and use of the infrastructure reasoning and interacting with software workshops through her/his own language. Therefore, the workshops act as cultural mediators among the different stakeholders by presenting the shared knowledge according to the language of each stakeholder.

3. Multimedia systems in the medical domain

The evolution of information technology may provide a valuable help in supporting physicians' daily tasks and, more importantly, in improving the quality of their medical diagnosis.

In current medical practice, physicians have the aid of different types of multimedia documents, such as laboratory examinations, X-rays, MRI (Magnetic Resonance Imaging), etc. Physicians with different specializations usually analyze such multimedia documents giving their own contribution to the medical diagnosis according to their "expertise".

However, this team of specialists cannot meet as frequently as needed to analyze all clinical cases, especially when they work in different hospitals or even in different towns or states. This difficulty can be overcome by providing physicians with computer systems through which they can cooperate at a distance in a synchronous and/or asynchronous way, also managing multimedia documents. In [7], we provide an example of such systems, that has been proposed to support neurologists working at the neurology department of the "Giovanni XXIII" Children Hospital of Bari, Italy, which gives them the possibility of organizing virtual meetings with neuro-radiologists and other experts, who may contribute to the definition of a proper diagnosis. The system is the result of an accurate user study, primarily aimed at understanding how the physicians collaborate in the analysis of clinical cases, so that functional and user requirements can be properly derived.

The study also revealed that physicians with different specializations adopt different languages to communicate among them and to annotate shared documents. For example, neurologists and neuro-radiologists represent two sub-communities of the physician community: they share patient-related data archives, some models for their interpretation, but they perform different tasks, analyze different multimedia documents (e.g., EEGs, in the case of neurologists, MRIs, in the case of neuro-radiologists) and annotate them with different notations, developed during years of experience. Such notations can be considered two (visual) languages.

The system described in [7] provides neurologists and neuro-radiologists with software environments and tools which are both usable and tailorable to their needs. It has been designed by adopting the SSW methodology [1]. Thus each specialist works with her/his own workshop to analyze the medical cases of different patients and to formulate her/his own diagnosis, taking into account the opinions of the other colleagues provided by the system, without the need of a synchronous consultation.

More specifically, if the neurologist needs to consult a neuro-radiologist, he makes a request by opening an annotation window. This window permits to articulate the annotation into two parts: the *question* to be asked to the colleague; and the *description* which summarizes information associated to the question. A third part can be tailored according to the addressee of the consultation request: if s/he is a physician who needs more details about the clinical case, the sender may activate the *detailed description* and fill it, otherwise s/he can hide it. In other words, the physician who wants to ask for a consultation is allowed to compose a tailored annotation specific to the physician s/he is consulting. In a similar way, a physician can make a different type of annotation in order to add a comment, which is stored and possibly viewed by other colleagues, thus updating the underlying knowledge base.

In the SSW approach, electronic annotation is a primitive operator, on which the communication among different experts is based. Moreover, the annotation is also a tool through which end users produce new content that enriches the underlying knowledge base. An expert has the possibility of performing annotations of various elements of the workshops, such as a piece of text, a portion of an image, a specific widget; through the annotation, the expert makes explicit her/his insights regarding a specific problem. The annotation is a peer-to-peer communication tool when it is used by experts to exchange annotated documents while performing a common task (e.g., defining a medical diagnosis). An expert can also annotate the workshop s/he is using, since annotation is also a tool used to communicate with the design team in charge of the maintenance of the system. The annotations are indexed as soon as they are created, by the use of a dictionary that is defined, updated and enriched by the experts themselves. The terms defined in the dictionary allow the experts to use the language, in which they are proficient, to annotate. They also permit the communication and understanding among the different actors having different expertise and languages.

4 EUD for managing Electronic Patient Records

The system described in the previous section allows its end users to perform some EUD activities. However, in the same medical domain, it is the management of the Electronic Patient Record (EPR) that pushes even more towards enabling EUD, as we will show in the following.

4.1 EPR

The current implementation of the EPR causes a lot of problems due to the fact that it is still commonplace that individual hospitals and even specific units within the same hospital, create their own standard procedures, so that physicians, nurses and other operators in the medical field are reluctant to accept a common unified format. Actually, they need to customize and adapt to their specific needs the patient record [8]. Thus, the EPR is a natural target for EUD.

Patient record is many-sided because it is a document to be read and understood by various and very different actors, such as physicians, nurses, patients' relatives, the family doctor, etc., so that it must have the ability to speak different "voices", i.e., to convey different meanings according to the actors using it [9].

The patient record contains at least two clear intertwined voices: a voice reporting what health professionals did to patients during their stay into the hospital; and another voice attesting that clinicians have honored claims for adequate medical care. Patient records are official, inscribed artifacts that practitioners write to preserve memory or knowledge of facts and events occurred in the hospital ward [10].

The patient record has two main roles: a short-term role refers to collect and memorize data to keep trace of the care during the patient's hospital stay; a long-term role refers to the archival of patient's data for research or statistical purposes [11]. Accordingly, the specialized literature distinguishes between primary and secondary purposes, respectively. Primary purposes regard the demands for autonomy and support of practitioners involved in the direct and daily care of patients; while secondary purposes are the main focus of hospital management, which pursue them for the sake of rationalizing care provision and enabling clinical research [9]. Our goal takes into account the primary purpose of patient record by designing an Electronic Patient Record (EPR) whose document structures and functionalities are aimed at supporting information inscription according to the specific needs of each involved stakeholder.

In this scenario, document templates and masks are usually imposed to practitioners, without considering the specific needs and habits of those who are actually using the EPR. The combination of requirements for both standardization and customization means that EPR systems are a natural target for EUD [9].

Again, in collaboration with the physicians of the "Giovanni XXIII" Children Hospital of Bari, Italy, we conducted a field study on the patient record and its use through unobtrusive observations in the wards, informal talks, individual interviews with key doctors and nurses, and open group discussions with ward practitioners. During the study, the analysts periodically observed the physicians during their daily work in the hospital (about 2-3 visits per month for two months). They observed how the identified stakeholders, i.e. head physicians, physicians, nurses, administrative staff, etc., of the same hospital manage paper-based patient records; our aim was to better understand which kind of documents, tools and languages are used. The information collected during the study has been used to identify the right requirements of an application implementing the EPR. The most important point that emerged is that they actually have specific patient records for each ward, even in the same hospital; this because there is the need of storing different data in the EPR, depending on the specific ward. For example, in a children neurological ward, information about newborn feeding must also be available, while in an adult neurological ward, information about alcohol and/or drug

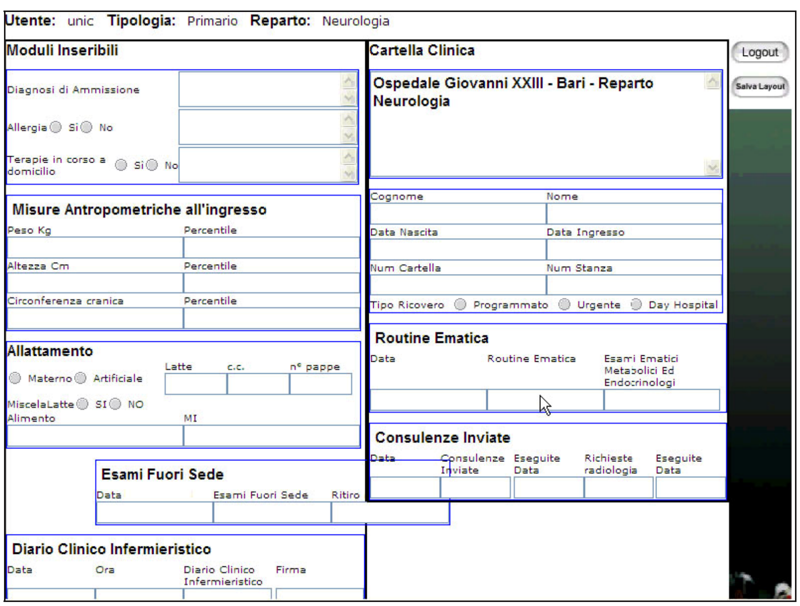


Figure 1. A screen shot of the SSW for the head physician "unic" of the "Neurologia" (neurology) ward.

assumption is required.

The different patient records can be seen as being composed by modules, each one containing specific fields for collecting patient data. Various stakeholders use the patient record in different ways and to accomplish different tasks, i.e., the nurse records the patient measurements, the reception staff records the patient personal data, the physician examines the record to formulate a diagnosis, and so on. We realized that the patient records used in different wards assemble a subset of modules in different ways, customized to the need of the specific ward. Thus, our approach was to identify the data modules that have to be managed in the whole hospital and let each head physician to design the EPR for her/his ward by composing a document through direct manipulation of such modules.

4.2 Co-designing the EPR with end users

The design of a prototype system to manage EPR followed the SSW approach, creating a software environment (SSW) for each type of stakeholder to allow them to accomplish their daily tasks in a comfortable and suitable way, as well as to give them the possibility of tailoring the SSW through EUD activities.

In particular, an SSW has been developed for the head physician, in which he can design the ERP tailored to the needs of his ward. The system supports his design activity by providing the SSW for the head physician with the set of predefined modules, among which he chooses those appropriate for his ward and assembles them in the layout he prefers. Figure 1 shows the SSW for the neurology head physician ("Primario Reparto: Neurologia" in Italian). The working area of the SSW is divided in two parts: on the left part there are all modules he can insert in the ERP ("Moduli Inseribili" in Italian), e.g., "Misure Antropometriche all'ingresso", "Allattamento", "Esami Fuori Sede", etc. ("Entrance Anthropometric Measurements", "Feeding", "External Examination" in English respectively),; on the right part there are the modules he is using to compose the tailored ERP ("Cartella Clinica" in Italian), e.g., personal data, "Routine Ematica" and "Consulenze Inviate" ("Hematic Routine" and "Sent Counsels" in English respectively). It does this by simple drag and drop of a module selected on the left part and inserting it in the desired position in the EPR he is composing in the right part of the working area.

Once the EPR design is completed, the head physician clicks on the "Save" button. In this simple way, he has actually created a software artefact that will be used by his ward personnel.

Figure 2 shows the EPR designed for neurology ward as it appears in the SSW for nurses. A nurse uses the EPR to primarily input data about patients. This end user does not have all EUD possibilities allowed to the head physician in his SSW, her/his tailoring is limited to modify the layout of the EPR modules. This because, if the nurse has to insert data in some specific modules, s/he prefers to move these modules to the top in order to find them quickly. Figure 2 shows a

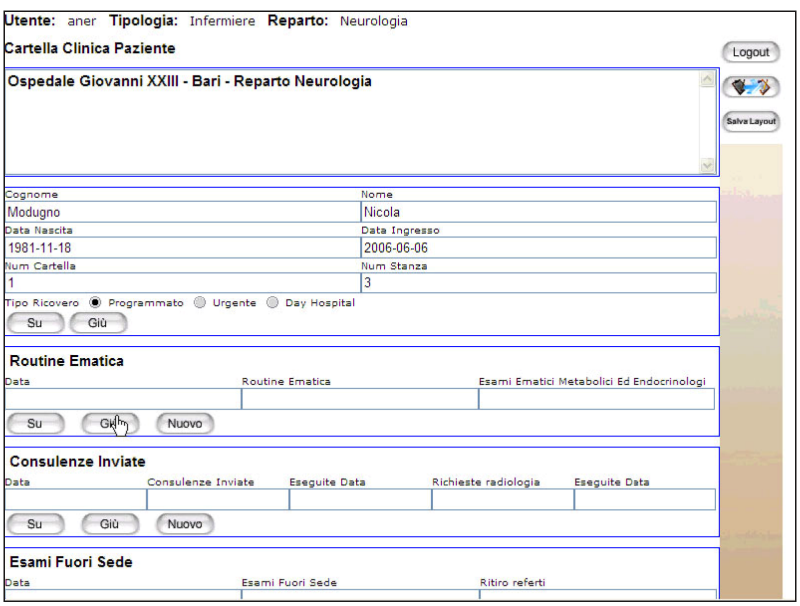


Figure 2. A screen shot of the SSW for the "Neurologia" (neurology) ward nurse "aner".

situation in which the pointer is on the module "Routine Ematica" ("Hematic Routine" in English) because the nurse wants to move this module in a different position.

5. Conclusions

This paper has discussed how to support end users who are increasingly willing to become co-designers of their tools and content. It is argued why End-User Development is particularly needed in the medical domain, were physicians, nurses, radiologists and other actors in the field are the end users. Furthermore, it is shown how the SSW methodology, which has been defined to create interactive systems that support EUD, can be successfully applied to this domain.
The infrastructure proposed by the SSW methodology to create interactive systems as a network of software environments (the SSWs) is implemented by exploiting a suite of XML-based languages. Specifically, the SSWs of the EPR prototype are implemented as IM^{2}L programs that are interpreted by a specialized engine, which is a plugin of the web browser [12, 13]. IM^{2}L (Interaction Multimodal Markup Language) is an XML-based language for the definition of software environments at an abstract level. In other words, environment elements and their behaviours are defined in a way independent by cultural and context-of-use characteristics; such characteristics are specified through other XML-based documents. The engine interprets these documents to instantiate the EPR SSWs, which are rendered by an SVG viewer under the coordination of the web browser [14]. As future work, we have planned an experiment with the end users. We will consider as quantitative metrics both the execution time of the assigned tasks and the errors made by the users. From a qualitative point of view, we will administer a post-experimental survey based on the SUS (System Usability Scale) method [15].

6. Acknowledgments

This work was supported by the Italian MIUR and by EU and Regione Puglia under grant DIPIS and by the 12-1-5244001-25009 FIRST grant of the University of Milan.

7. References

1. M.F. Costabile, D. Fogli, P. Mussio and A. Piccinno, "Visual Interactive Systems for End-User Development: a Model-based Design Methodology," *IEEE Transactions on System Man and Cybernetics Part A-Systems and Humans*, vol. 37, no. 6, 2007, pp. 1029-1046.
2. G. Fischer and E. Giaccardi, "Meta-Design: A Framework for the Future of End User Development," *End User Development*, H. Lieberman, F. Paternò and V. Wulf, eds., Springer, 2006, pp. 427-457.
3. D.A. Schön, *The Reflective Practitioner: How professionals think in action*, Basic Books, 1983, p. 374.
4. A. Cypher, ed., Watch what I do: programming by demonstration, MIT Press, 1993.
5. P. Bottoni, M.F. Costabile and P. Mussio, "Specification and dialogue control of visual interaction through visual rewriting systems," *ACM Transactions on Programming Languages and Systems (TOPLAS)*, vol. 21, no. 6, 1999, pp. 1077-1136.
6. M. Costabile, P. Mussio, L. Parasiliti Provenza and A. Piccinno, "Supporting End Users to Be Co-designers of Their Tools," *End-User Development*, V. Pipek, M. B. Rosson, B. de Ruyter and V. Wulf, eds., Springer, 2009, pp. 70-85.
7. M.F. Costabile, D. Fogli, R. Lanzilotti, P. Mussio and A. Piccinno, "Supporting Work Practice Through End-User Development Environments," *Journal of Organizational and End User Computing*, vol. 18, no. 4, 2006, pp. 43-65.
8. C. Morrison and A. Blackwell, "Observing End-User Customisation of Electronic Patient Records," *End-User Development*, V. Pipek, M. B. Rosson, B. de Ruyter and V. Wulf, eds., Springer, 2009, pp. 275-284.
9. F. Cabitza and C. Simone, "LWOAD: A Specification Language to Enable the End-User Develoment of Coordinative Functionalities," *End-User Development*, V. Pipek, M. B. Rosson, B. de Ruyter and V. Wulf, eds., Springer, 2009, pp. 146-165.
10. M. Berg, "Accumulating and Coordinating: Occasions for Information Technologies in Medical Work," *Computer Supported Cooperative Work (CSCW)*, vol. 8, no. 4, 1999, pp. 373-401.
11. G. Fitzpatrick, "Integrated care and the working record," *Health Informatics Journal*, vol. 10, no. 4, 2004, pp. 291-302.
12. B.R. Barricelli, A. Marcante, P. Mussio, L. Parasiliti Provenza, M. Padula and P.L. Scala, "Designing Pervasive and Multimodal Interactive Systems: An Approach Built on the Field," *Handbook of Research on Multimodal Human Computer Interaction and Pervasive Services: Evolutionary Techniques for Improving Accessibility*, P. Grifoni, ed., Idea Group Inc., to appear.
13. D. Fogli, G. Fresta, A. Marcante and P. Mussio, "IM2L: A User Interface Description Language Supporting Electronic Annotation," *Proc. Workshop on Developing User Interface with XML: Advances on User Interface Description Languages, AVI 2004*, 2004, pp. 135-142.
14. W3C, "Scalable Vector Graphics (SVG)," 2009; http://www.w3.org/Graphics/SVG/.
15. J. Brooke, P.W. Jordan, B. Weerdmeester, A. Thomas and I.L. McLelland, "SUS: A quick and dirty usability scale," *Usability evaluation in industry*, Taylor and Francis, 1996.

Transformation from Web PSM to Code

Yen-Chieh Huang[1,2], Chih-Ping Chu[1], Zhu-An Lin[1], Michael Matuschek[3]
[1]*Department of Computer Science and Information Engineering, National Cheng-Kung University, Tainan, Taiwan*
[2]*Department of Information Management, Meiho Institute of Technology, Pingtung, Taiwan*
[3]*Department of Computer Science, University of Duesseldorf, Germany*
E-mail :p7894121@mail.ncku.edu.tw

Abstract

*This research proposes how class diagrams that use the Unified Modeling Language **(UML)** can be converted to a user interface of a Web page using the Model Driven Architecture **(MDA).** From the Platform Independent Model **(PIM)** we go to the Web Platform Specific Model **(PSM)**, and then to the direct generation of code templates for Web page applications. In this research the class diagrams are drawn with the Rational Rose, then, using our self-developed program, these diagrams can be transformed into code templates with Servlets, JSP, and JAVA. We implement a case study for verification, and then calculate the transformation rate with lines of code **(LOC)** coverage rate by measuring the LOC after transforming and after the system is finished. The results show the transformation rate is about thirty-six to fifty percent, which represents that this research can help the programmers to greatly reduce the developing period.*

Keywords: Model Driven Architecture, Platform Independent Model, Platform Specific Model

1、Introduction

Software is largely intangible [1]. Software development gradually transforms from structure analysis and design to object-oriented analysis and design, but the software industry is labor intensive, even after finishing system analysis, the programmers still start from scratch and write the code. Especially in the application software development for Web pages, in the last few years, there are many researches have been proposed to reduce code and development time. This research focuses on how class diagrams can be transformed into Web pages, the results could reduce the development time for Web pages programmers. The common Web pages developing tools include JSP, PHP, and ASP etc.. The platform used in this research is JAVA, the Web pages developing tool is JSP, relevant technology are JSP, Servlets and Ajax. This research uses IBM Rational Rose as the CASE tool for class diagram object modeling, and the user interface code templates are then created via the conversion program written by ourselves.

2、Literature Review

The object-oriented paradigm has gained popularity in various guises not only in programming languages, but also in user interfaces, operating systems, databases, and other areas [2]. Classification, object identity, inheritance, encapsulation, and polymorphism and overload are the most prominent concepts of object-oriented systems [3]. The UML is a modeling language that helps describing and designing software systems, particularly software systems built using the object-oriented approach. This research uses Robustness diagrams [4] for describing the application environment of Web pages.

The MDA is a framework for software development defined by the Object Management Group (OMG). It is the importance of models in the software development process [5, 6]. The MDA development life cycle included four kinds of models. **Computation Independent Models** (CIM) describe the requirements for the system and represent the highest-level business model. It is sometimes called "domain model" or "business model". A **PIM** describes a system without any knowledge of the final implementation platform, and this PIM is transformed into one or more PSMs. A **PSM** is tailored to specify a system in terms of the implementation constructs that are available in one specific implementation technology. The final step in the development is the transformation of each PSM to **code**. The **CIM, PIM, PSM**, and **code** are shown as artifacts of different steps in the software development life cycle, which is shown in Figure 1.

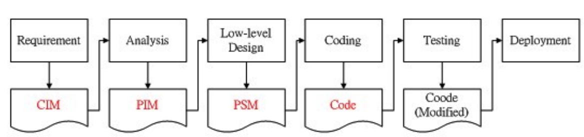


Figure 1. MDA software development life cycle and output artifacts

The most widely used architecture in the environment of Web applications is Browser/Server (B/S) approach, an example for a specific Client/Server (C/S) structure [7]. The basic architecture of Web systems includes a client browser, a Web server, and a connecting network. The principal protocol for communication is the **Hypertext Transfer Protocol (HTTP)**. The principal language for expressing the context between the client and the server is **Hypertext Markup Language (HTML)** [8].

Relevant technologies for today's Web applications include CGI, Applets, ActiveX controls, plug-ins and Ajax etc. To explain the general structure of such a Client/Server system, a Web page can be modeled into a class, and a client page can be modeled into another class, which must be drawn by the method of extending UML [9].

3、Transformation from Class Diagrams to Web Applications

In the concept of MDA we must first create the PSM design for a specific Web application. A Web page can be expressed by class diagrams where every stereotype (including stereotype classes and associations) is defined in order to describe the situation of every Web page, then the Web class diagrams can be drawn and, in the final step, it can be transformed into a code template.

3.1 Web Pages Components Mapping Methods

3.1.1 Stereotypes

In order to extend its function of use in UML, we can use stereotypes to strengthen and define the class model. Stereotypes allow us to get a more proper description to the class objects, they can be used for describing and limiting the characteristics of the module components, and they exist in standard UML components [10]. In this paper, we use Rational Rose to define control classes and strengthen the classes that describe the Web pages. This research proposes stereotype class mapping methods as described in Table 1.

3.1.2 Association Stereotypes

In order to implement Web modules, it is vital to control user-site and server-site requests and responses via HTML in the network. Using association stereotypes between classes is an optional way to model HTTP parameters, and it is useful when parameters are relatively complex or have special semantics and extra documentation is necessary. Therefore, this research proposes the mapping methods of association stereotypes between classes as shown in Table 2.

3.2 PSM to Code Template Transformation

Every stereotype class has different transformation model, in here; we describe a Servlet transformation rule as an example. The attributes and methods in Servlet are implemented by traditional JAVA, but the difference lies in the association between classes. Generally speaking, a Servlet must

Table 1. Stereotypes Mapping in Class

Stereotypes	Description
<<Servlet>>	Responsible for showing the request of client site, and communicating with back end module The methods of this class contain at least Get() or Post().
<<Server Page>>	A server page represents the server site information, the attributes and methods in this class are implemented by Scripting Element.
<<Client Page>>	A client page represents the <HTML> element, which has two principal child elements: <HEAD> and <BODY>. The <HEAD> represents structural information about the Webpage; the <BODY> element represents the majority of the displayed content [8].
<<Form>>	The HTML <<Form>> stereotype class represents some attributes, such as input boxes, text areas, radio buttons, check boxes, and hidden fields, these classes map directly to a <Form> element [8].
<<Model>>	A <<model>> stereotype class represents the logical operation of business processes, which is implemented by JAVA. Its meaning is the same as traditional class diagrams, therefore a class diagram notation can ignore the <<Model>> stereotype in this research.

Table 2. Association Stereotypes

Association	Description
<<Build>>	This is an action of a Servlet or a Server Page creates a Client Page or a Form.
<<Link>> [8]	A relationship between a client page and a server-side resource or Web page.
<<Include>>[8]	A directional association from a Web page to another Web page.
<<Redirect>>	The client page should be automatically replaced with another client page, where Post and Get are two methods to achieve this, among others.
<<Object>> [8]	This represents many types of embedded objects, such as Applet, ActiveX controls. The parameters for the object are defined in the parameterized class.
<<Asynchronous>>	The client page sends an asynchronous request to Servlet.
<<Submit>>	A relationship between a <<Form>> and a server page. Post or Get are used for submitting, among other methods.

accept a Form request, and then a redirection to another Webpage occurs. Its transformation steps are as follow:

1. <<Form>> request- According to Form request the association names (Get or Post), then declare the method of doGet or doPost.
2. <<Client Page>> asynchronous- In Servlet, implement the asynchronous pattern and then declare the method doAsynWork.
3. << Redirect >>- Generate the code as follow:

```
RequestDispatcher view =
request.getRequestDispatcher("/****Redirect Page
***/");
view.forward(request, response);
```

4、Measurement

For the experimental evaluation we adopt "code coverage" to calculate the result. Code coverage is a measure used in software testing. It describes the degree to which the source code of a program has been tested. In this research, code coverage represents the ratio of information in class diagrams to the information in the full implemented system. Talking about information, we define the way of measurement and standard of quantification analysis as follows:

4.1 The Way of Measurement

In a software development project, software measurement can be achieved in a lot of ways, such as lines of code (LOC), function point (FP), object point, COCOMO, and Function requirement etc. We choose LOC, and the reasons are:

1. The value is easily measured.
2. There is a direct relationship to the measurement of person-months (effort).
3. Effort is also a size-oriented software metric [11].

For a class diagram, it expresses static information as well as the relation between classes, and the resulting LOC can be easily counted automatically after transformation.

4.2 Counting Standard

LOC counters can be designed to count physical lines, logical lines, or source lines by using a coding standard and a physical LOC counter. For different kinds of Coding Style, the LOC turns out differently, so we need to define the Coding Standard and Counting Standard which we use for our measurement.

In this research, line counters are defined as follows:

1. XML has defined and self-defined tags in Web pages, a set of tag counts as one line.
2. If the web pages are not XML, (e.g. Scripts, Scriptlets, and Expressions), every line of code counts as one line.

5、 Case Study

5.1 Experiment Steps

The CASE Tool selected for this experiment is the Rational Rose from IBM which transforms class diagrams into code templates. First, Rational Rose is used to draw the class diagrams, then the labels of the stereotypes are added in the class diagrams, and lastly we utilize the program developed by ourselves to transform the class diagrams into code templates.

5.2 Case Description

To verify the theoretical structure proposed by this research we use the practical example of a Login/Register System. It has three main functions in the Use Case Diagram. There are "Account registration", "User login", and "Display Home page".

Figure 3 is a class diagram of PIM of a user Login/Register System which reflects the Use Case diagrams. In the preliminary design, which uses Robustness diagrams for description, we include the entity classes, boundary classes and control classes. Boundary classes represent the shown Web page content, i.e. the information in the system, such as the account and password fields that LoginClient offers for the user login. Control classes deal with the parameter request by the boundary classes, such as login request to LoginServlet of LoginClient, and they are determined to call out Register of the Entity class to deal with the request.

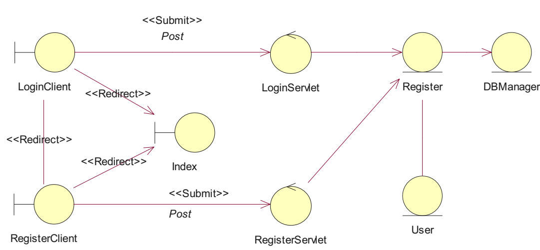


Figure 3. The PIM of a Login/Register System

Use Case 1: Account Registration

This use case includes the boundary classes RegisterClient, RegisterForm, and RegisterBackForm, the control class RegisterServlet, and the entity class Register as back end. Between the classes RegisterClient and RegisterServlet, there are asynchronous relations, so the Ajax pattern will be used for realizing the code transformation.When the user succeeds to register, the class RegisterServlet will redirect him to the Index home page.

Use Case 2: User Login

This use case includes the boundary classes LoginClient, LoginForm, and LoginToRegister and the control class LoginServlet. When the user inputs his account and password, the class LoginForm will send a request to the class LoginServlet using the Post method, and then the class LoginServlet makes the decision if the user is redirected to the Index or the class LoginClient.

Use Case 3: Display the home page

The home page includes Index and AVLTreeApplet, and is displayed by a Java Applet. It is described how the Applet object is loaded and integrated into the Index home page via object parameter classes.

5.3 Measurement Result

We measured the LOC of the code template for each use case after transformation and the LOC of the finished system by the previously defined counting standard. The data is shown in Table 3.

Table 3. Measurement Result

	LOC of Code template after transform	LOC of finished system	Transform ratio
Use Case 1: Account Registration			
registerclient.html	42	95	44%
RegisterServlet.java	14	38	37%
Use Case 2: User Login			
loginclient.html	11	22	50%
LoginServlet.java	11	21	50%
Use Case 3: Display home page			
index.jsp	5	14	36%

The results show that the transformation rate is about thirty-six to fifty percent. When we focus on the part not responsible for the program logic in this class, this is a relatively high proportion. The transformation into the code template according to the defined Web page class diagrams represents the static structure model of the system, consisting of attributes, operations, and associations between classes. However, the system operation logic cannot be expressed in detail. This part is still up to the programmers.

6、 Conclusion

Nowadays, Web code must be programmed from scratch even if the PSM analysis is finished, but in this research we proposed a method of code template transformation. By adding stereotypes to class diagrams, they can describe Web pages, synchronous or asynchronous relations, and we can transform them into code templates with distinct logical, control, and view code blocks using JSP&Servlets or the MVC model.

Asynchronous relations can be realized using many methods. This research adopts Foundations of Ajax to express that the client site responds to the server site. Furthermore, reverse engineering is a factor to be considered, so that maybe change of the code can be reflected in the Web class diagrams afterwards.

For the case study example in this research, a class diagram transformed into code templates can only be achieved about thirty-six to fifty percent of the whole system, which expresses that it does not discover all the sufficient information we want.

Because of the definition of class diagrams and the representing models, they can only express static class content and relationships. There are also other aspects that cannot be described in design and transformation for more complicated program logic.

For this reason, we can make use of sequence diagrams, and state diagrams, in order to describe the dynamic call and transfer between the states. So, the further research will study how to create Web code templates from interaction diagrams and behavior diagrams.

Reference

[1] Lethbridge, T.C. and Laganiere, R., *Object-Oriented Software Engineering: Practical Software Development using UML and JAVA Second Edition*, Mcgraw-Hill, 2005.

[2] Nierstrasz, O., *A Survey of Object-Oriented Concepts, In Object-Oriented Concepts, Databases and Applications*, W. Kim and F. Lochovsky, , ACM Press and Addison-Wesley, 1989, pp. 3-21.

[3] Gottlob, G., Schrefl, M. and Rock, B., Extending Object-Oriented Systems with Roles, *ACM Transactions on Information Systems*, Vol. 14, No. 3, Jul. 1996, pp. 268-296.

[4] Ambler, S.W., *The Object Primer: Agile Model-Driven Development with UML 2.0*, Cambridge Univ Pro, 2004.

[5] Kleppe, A., Warmer, J. and Bast, W., *MDA Explained: The Model Driven Architecture™: Practice and Promise*, Addison Wesley, Apr. 2003.

[6] Koch, N., Classification of model transformation techniques used in UML-based Web engineering, *IET Software*, Vol. 1, Issue 3, Jun. 2007, pp. 98-111.

[7] Li, J., Chen, J. and Chen, P., Modeling Web Application Architecture with UML, *IEEE CHF*, 30 Oct. 2000, pp. 265-274.

[8] Conallen, J., *Building Web Applications with UML Second Edition*, Addison Wesley, 2002.

[9] Conallen, J., Modeling Web Application Architectures with UML, *Communications of the ACM*, Vol. 42, No. 10, Oct. 1999.

[10] Djemaa, R.B., Amous, I., and Hamadou, A.B., WA-UML: Towards a UML extension for modeling Adaptive Web Applications, *Eighth IEEE International Symposium on Web Site Evolution*, 2006, pp. 111-117.

[11] Humphrey, W.S., *PSP A Self-Improvement Process for Software Engineers*, Addison Wesley, Mar. 2005.

Experiences with visual programming in Engineering Applications

Valentin Plenk
University of Applied Sciences Hof
Alfons Goppel Platz 1, 95028 Hof, Germany
valentin.plenk@fh-hof.de

Abstract

In children's playrooms and in secondary school projects programmable toys and entry level programming courses use visual programming languages instead of the (standard) textual source codes seen in Logo, BASIC, Java. Higher education and research also propose visual programming or even (graphical) model based design to steepen the learning curve[1].

Industry however appears unfazed with this approach. Textual source code is still the main means of representing software.

Based on experience gained in laboratory exercises conducted with students of an undergraduate course in mechatronics this paper adresses the feasibility and efficiency of this approach.

1. Introduction

A wide range of research papers proposes graphical representations for complex software ranging from domain specific code generators (e.g:[6], [9]) to software models expressed in UML (e.g:[4], [18]).

[3] succinctly summarizes the reasoning for the visual representation :

> The human visual system and human visual information processing are clearly optimized for multi-dimensional data. Graphical programming uses information in a format that is closer to the user's mental representations of problems, and allows data to be manipulated in a format closer to the way objects are manipulated in the real world.
>
> Another motivation for using graphics is that it tends to be a higher-level description of the desired action (often de-emphasizing issues of syntax and providing a higher level of abstraction) and may therefore make the programming task easier even for professional programmers.

This research effort is flanked by a wide range of commercially available, domain specific visual programming and execution environments. Some examples could be National Instrument's Labview, Agilent's Vee, IEC 61131-3 Sequential Function Charts, Mathworks's Simulink. The following links on Wikipedia give a quick synopsis of these products: [12, 13, 14, 15, 16, 17]. More information can be found at the respective products' websites.

In daily practice the unified modeling language [7] has become a (graphical) standard in the early phases of the software development process – i.e: in design documents. The diagrams are used to describe the architecture of a software product on a more or less abstract level. Recently efforts to execute the models have become visible.

However the vast majority of actual software products is still implemented in textual source code. Common sense apparently considers the available tools as unprofessional or unsuited for big projects. There are little to none publications investigating the validity of this opinion.

To contribute some facts this paper summarizes experiences with visual programming made in an undergraduate course at the University of Applied Sciences Hof. A group of engineering students specializing in mechatronics was tasked with a signal-processing exercise.

2. Mechatronics

The students in the bachelor course "industrial engineering" specializing in mechatronics have to master a series of laboratory exercises designed to deepen the understanding of signal processing theory and its application in mechatronical systems. In one of these exercises the students are tasked with defining and implementing a criterion for stopping the motor of a car's power-window when something or

[1] In this context a "steep" learning curve means quick progress in learning – the increase in knowledge over time is growing steeply (at least during the initial stages of learning).

someone is clamped in the window.

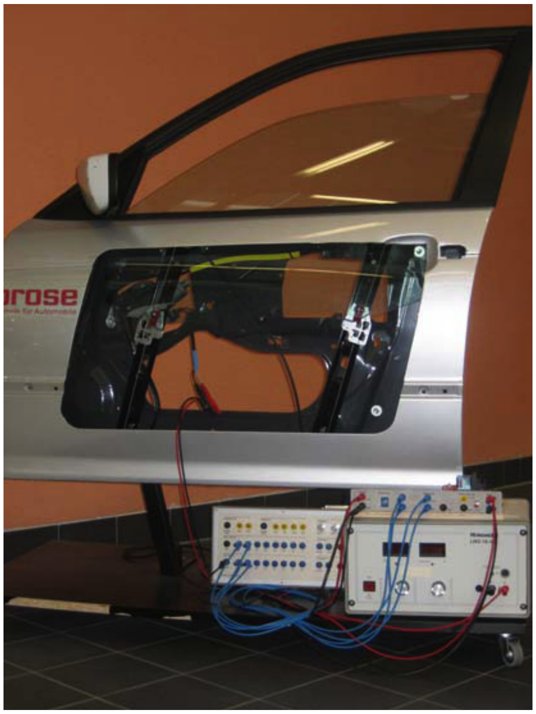


Figure 1. Experimental setup for the mechatronical assignment

2.1. The laboratory exercise

The laboratory setup comprises a car-door with a power-window, power electronics to drive the mechanics, measurement circuitry to pick up the motor current and a controlling PC running Agilent Vee (figure 1).

"Agilent VEE is an easy-to-use intuitive graphical test & measurement software that provides a quick and easy path to measurement and analysis." [1] The software offers visual dataflow programming. The students "only" need to connect a few blocks – signal source, signal processing, signal sink – to implement their first application, a simple data-acquisition and display program necessary to analyze the motor current data to find criteria for detecting that something is clamped in the window. Figure 2 shows a solution for this first step.

This first application's main function is the dataflow from the analogue-digital-converter (data source block at the right) to the scope display (sink block at the far right), represented by the thin line linking source and sink. The remaining blocks to the left of figure 2 are necessary to setup the ADC and to implement an infinite loop to read and display the data. This program sequence is specified with a second type of connecting lines representing the control flow (thick lines).

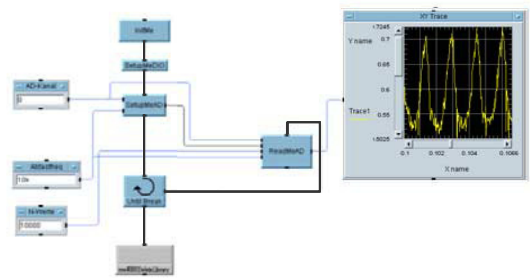


Figure 2. Simple program to study the behaviour of the motor current

The students are then instructed to conduct a series of experiments with and without objects clamped in the window to find a criterion for detecting that something is clamped in the window. This step profits from the intuitive way of combining signal-sources, signal-processing-blocks and sinks (displays...) offered by the dataflow-oriented software.

Once a criterion is established, its implementation which is usually a combination of calls to existing signal-processing-blocks is quickly found – again thanks to the dataflow design.

Making the criterion stop the motor however is not so easy. The intuitive program used so far is straightforwardly extended to periodically read the ADC and the window-buttons and to write the motor up/down bits. The criterion still works, since it is connected to the dataflow from the ADC. As soon as the criterion stops the motor, the data from the ADC will no longer indicate a clamped object (the motor current will be zero), thus the criterion will allow the motor to run (again). This will lead to data that make the criterion stop the motor, which makes the criterion start the motor. Eventually the window will perform a jerking motion.

To overcome this problem the students have to add states to the software. They can do it in a dataflow compatible way, by adding a feedback variable that is written at the end of the dataflow and read at the beginning of the dataflow. The students usually reject this approach as unintuitive. The alternative approach is a more complex controlflow consisting of two nested loops: The outer loop is the data-acquistion loop used so far. The inner loop is entered once the criterion was activated and blocks the application until the window-buttons are released.

Figure 3 shows a rather well structured solution for the exercise. The software is still functional but no longer easy to understand. The original loop is marked with "data acquisition". The additional code is necessary to implement the criterion and the "unjerky" stop function.

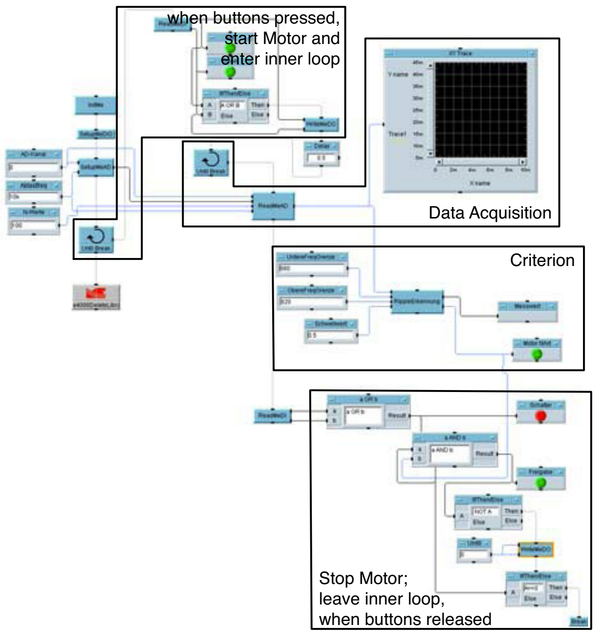


Figure 3. Example for a (good) solution

2.2. The learning curve

Section 2.1 gives an impression of the exercise's complexity and size. The exercise is run in a 4 hour session with one experienced tutor for four to five groups. Each group with four to five students works in parallel on different exercises. The students have no prior knowledge of Agilent Vee. The instructions for the exercise contain some hints regarding the criterion and the usage of Agilent Vee. The initial program shown in figure 2 is a part of these instructions.

Good students require about 1 hour of tutoring with respect to the criterion and Agilent Vee and are then able to implement a working software like the version shown in figure 3.

This astounding performance is attributed to the visual programming interface offered by Agilent Vee. Other exercises of similar complexity implemented in C++ or BASIC show three to four times lower productivity even though they are run with undergraduate students in computer science that have extensive prior knowledge of the programming language and the environment.

The key element seems to be Agilent Vee's intuitive way to connect (existing) software blocks by dataflow lines. In C++ or BASIC such a link would be implemented as a function/method-call and probably require some data conversion from one call to the next call. Agilent Vee handles the conversions transparently and therefore allows the user to concentrate on the application.

But there is no rose without a thorn – the initially extremely steep learning curve in this exercise becomes considerably flatter as soon as the students need to add the control flow mechanisms to prevent the jerky movement. At this point they ask the tutor to find an implementation for their solution.

2.3. Evolvability

Figure 3 shows that the initially attractive, intuitive visual programming quickly becomes a poorly structured confusing diagram of linked blocks. Without the manually inserted boxes, the code is almost unintelligible even though this example is fairly simple. More complex applications will have an even more confusing structure and therefore need even more documentation.

The necessity to use structuring elements – e.g: hierarchical blocks – along with the increased need for documentation significantly reduces the productivity of the visual approach.

In "one-shot" projects where no revisions are necessary this lack of evolvability is not a problem. [10] describes a field of industry using almost only this kind of software projects. In this context the main challenge is the reuse of code modules in a large software framework. With the right kind of "building blocks" in the visual programming environment the reuse of powerful code modules is facilitated greatly.

3. Conclusion

The results clearly show that with visual programming the learning curve is indeed steep compared to textual source codes. The students produced impressive results rather quickly, especially as long as big code-blocks are reused by coupling them together to calculate the criterion. To stop the motor properly control flow elements have to be used as well. This rapidly results in a complex diagram, that might be hard to evolve in future versions.

In the authors' opinion visual programming is a powerful approach that allows to quickly build highly functional applications that efficiently reuse code-blocks. On the other hand these applications are not evolvable and should be considered as "one-shot" customization requiring a complete (quick) rewrite for the next version. [2, 5, 6, 8, 9, 18] exemplify the demand for this kind of software projects / tools in industry.

References

[1] Agilent. Agilent VEE Pro 9.0. http://www.home.agilent.com/agilent/product.jspx?cc=US&lc=eng&ckey=1476554&nid=-34448.806312.00&id=1476554&cmpid=20604, February 2009.

[2] M. C. Andrade, C. E. Moron, and J. H. Saito. Reconfigurable system with virtuoso real-time kernel and tev environment. *Symposium on Computer Architecture and High Performance Computing*, pages 177–184, 2006.

[3] C. Andronic, D. Ionescu, and D. Goodenough. Automatic code-generation in a visual-programming environment. *Proceedings of the Canadian Conference on Electrical and Computer Engineering*, pages 6.30.1–6.30.4, September 1992.

[4] J. Bartholdt, R. Oberhauser, and A. Rytina. An approach to addressing entity model variability within software product lines. *3rd International Conference on Software Engineering Advances*, pages 465–471, 2008.

[5] G. Bayrak, F. Abrishamchian, and B. Vogel-Heuser. Effiziente Steuerungsprogrammierung durch automatische Modelltransformation von Matlab/Simulink/Stateflow nach iec-61131-3. *Automatisierungstechnische Praxis (atp)*, 50(12):49 – 55, December 2008.

[6] J. C. Galicia and F. R. M. Garcia. Automatic generation of concurrent distributed systems based on object-oriented approach. *Proceedings of the IASTED International Conference on Parallel and Distributed Computing and Systems*, 16:809 – 814, 2004.

[7] O. M. Group. Uml standard. http://www.omg.org/spec/UML/2.1.2/, February 2009.

[8] R. Härtel and T. Gedenk. Design and implementation of canopen devices. *Proceedings of the iCC2003*, pages 04–1 – 04–6, 2003.

[9] P. Lauer. Auto generated production code for hydraulic systems. *Proceedings of the 6th International Fluid Power Conference, Dresden*, 1:171 – 182, 2008.

[10] V. Plenk. A benchmark for embedded software processes used by special-purpose machine manufacturers. In *Proceedings of the third international Conference on Software Engineering Advances*, pages 166 – 171, Los Alamonitos, 2008. CPS.

[11] P. Stöhr, S. Dörnhöfer, R. Eichinger, E. Frank, D. Krögel, M. Markl, M. Nowack, T. Sauerwein, J. Steffenhagen, and C. Troger. Untersuchung zum Einsatz von Robotern in der Informatikausbildung. internal report University of Applied Sciences Hof, July 2008.

[12] Wikipedia. Agilent Vee. http://en.wikipedia.org/wiki/Agilent_VEE, February 2009.

[13] Wikipedia. Grafcet. http://de.wikipedia.org/wiki/Grafcet, February 2009.

[14] Wikipedia. Labview. http://en.wikipedia.org/wiki/Labview, February 2009.

[15] Wikipedia. Mindstorms NXT. http://en.wikipedia.org/wiki/Mindstorms_NXT, February 2009.

[16] Wikipedia. Sequential function chart. http://en.wikipedia.org/wiki/Sequential_function_chart, February 2009.

[17] Wikipedia. Simulink. http://en.wikipedia.org/wiki/Simulink, February 2009.

[18] D. Witsch, A. Wannagat, and B. Vogel-Heuser. Entwurf wiederverwendbarer Steuerungssoftware mit Objektorientierung und UML. *Automatisierungstechnische Praxis (atp)*, 50(5):54 – 60, 2008.

Advantages and Limits of Diagramming

Jaroslav Král Michal Žemlička
Charles University in Prague, Faculty of Mathematics and Physics
Department of Software Engineering
Malostranské nám. 25, 118 00 Praha 1, Czech Republic
{Jaroslav.Kral,Michal.Zemlicka}@mff.cuni.cz

Abstract

The importance of software diagrams is often overemphasized as well as underrated. We show that diagrams must be used especially in the cases when (weakly structured) texts related to the natural language must be also used. Examples are requirements specification, software architecture overviews, formulation of ideas or basic principles of solutions. The fact that texts must be used together with diagrams is often considered as a disadvantage as both "formats" must be synchronized what is not an easy task. We show that a proper combination of advantages and disadvantages of texts and diagrams can bring a great benefit.

The advantages of diagrams in late stages of SW development are not clear. One solution is to use diagrams in initial stages of development only. An example of such strategy is agile development. The second way is to suppress the importance of code like in model-driven architecture (MDA). MDA works well for small and routine projects. The application of diagrams in large projects leads to complex systems of complex diagrams. Such diagrams are not too useful. It may be the main reason of limited success of MDA.

1. Introduction

Diagrams and the practices using them are considered to be very helpful, easy to understand and use. Experience indicates that the use of diagrams is not without issues. The notations of diagrams have been evolving quite rapidly, may be quicker than software development paradigms. Some aspects of software system structure and use have several different modeling frameworks. There are several diagram types used to model the same entities, e.g. business processes or workflows. Workflow can be described by activity diagrams in UML, diagrams in Aris [4], there are also two system for workflow model languages designed by W3C and WfMC. It indicates that the modeling needs and the properties of best modeling practices are not clear enough. The semantics of the diagrams is vague. Under certain circumstances it need not be wrong. The diagrams do not support newly invented constructs – an example is service government (compare the history of exceptions in flow charts).

There are doubts whether diagrams are of any use in software maintenance as the updates of code and updates of diagrams are usually not well synchronized and the diagrams therefore tend to be obsolete. Some methodologies like extreme programming [2] forbid any use of diagrams for maintenance or require, like in Agile Programming Manifesto [3] that the diagram should be used as an auxiliary mean only. An intended exception is Model Driven Architecture (MDA, [7]) when code has an auxiliary role and it is generated from diagrams. It has some drawbacks discussed below.

On the other hand the use of diagrams in early phases of development is quite common. But it can, as noticed, lead to the situation when a software system has two defining/describing documentations – code and supporting diagrams being often obsolete.

2. Engineering Properties of Diagrams

The graphical nature of diagrams implies the following drawbacks

1. The diagrams consisting of many entities are unclear as humans are unable to follow more than ten entities at once. Diagrams therefore are not too advantageous to model complex systems. It is confirmed by observation. The solution can use decomposition of the systems into autonomous components (e.g. services in SOA) and hierarchical decomposition using subdiagrams. The subdiagrams depict subsystems. The problem is that it is often difficult to do it well technically and conceptually.

2. It is often difficult to implement a "good" modification of diagrams, i.e. transformations retaining desira-

ble properties of transformed diagrams like lucidity. It implies that the use of diagrams during software maintenance need not be helpful.

3. The semantics of diagrams tends to be vague in order to support intuitiveness and flexibility. It is good for specification as in this case the semantics can be gradually ”tuned”. It, however, partly disqualifies the diagrams as a code definition tool.

All the facts are straightforward. Their managerial consequences for process control are often not properly taken into account. The consequence is that diagrams tend not to be useful for the maintenance of long living systems.

3. Diagrams in Early Stages of Software Development

Diagrams are used in early stages of software development. They are often used in requirement specification documents (RSD). RSD can be and MDA must be highly formalized but it then needs not be any good solution as the semantics of such formalized specification language is rather IT knowledge domain oriented than user knowledge domain oriented. It can disturb the focus on the user visions and user needs as the semantics of RSD can far from the semantics of the user-domain languages; It therefore almost avoid the possibility of effective collaboration with users during the formulation of the specification[1].

A satisfactory solution is to use a specification language close to the professional user knowledge domain language [5] and to use user-domain diagrams. The diagrams like the specification languages should be flexible enough to enable iterative specification techniques and stepwise increase the precision and depth of requirements.

Such diagrams are then well understood by users so they can well collaborate with developers. In this case the user knowledge domain diagrams can be and usually should be used. Such diagrams are used as long as the specification documents are used and updated.

If a larger system is to be developed, its overall architecture must be specified together with the requirements specification as the architecture determines the structure of the requirements specification document. It is particularly typical for systems having service-oriented architecture (SOA). The diagrams depicting some aspects of SOA are very useful. Other aspects are difficult to depict yet.

It is often preferable to depict other overall (global) properties of the solution. The proper use of diagrams can substantially speed up the specification process and enhance the quality of the resulting specification.

[1] It is one of the reasons why MDA has only a limited success.

As the specification is a crucial document, sometimes even a part of formal agreements, it is kept actual and the above problems with obsoleting diagrams need not take place. Diagrams can help to explain global properties of systems.

A crucial fact is that the diagrams are associated with text in a ”natural” language – requirements in the form legible for customers, informal descriptions of system architecture or of some aspects of the solutions depicted by the diagrams.

3.1. Why Diagrams?

Diagrams can be something like a ”materialization” of ideas. They, like any natural language, can be as vague or incomplete as necessary at a given moment or according ”state of art” of a project. They can hide details but they can be iteratively precised to achieve needed exactness and completeness. It is simplified by the fact that they can be well integrated into text documents.

Many diagram types are intuitive and are the part of professional languages. They should increase transparency what is possible if they are not too complex, otherwise they can be worse than a structured text.

Some diagramming techniques provide an excellent tool for thinking and enable an easy detection of thinking gaps.

The applications or the use of diagrams in specification documents increased the legibility and ”visibility” of the requirements and supports the collaboration of developers and users. It is very important as the snags in specifications causes 80 % of development failures.

Diagrams are intuitively easier to understand by both developers and users. Almost no tiresome preliminary training of users e.g. the reading of manuals and syntax training is necessary. Diagrams are part of many user knowledge domains. And as such they can be used in specification documents.

Some global properties like the system architecture are well depicted by proper diagrams. Incomplete diagrams can be useful. Iterative development of diagrams supports an iterative thinking as a multistep approximation process.

The missing or incomplete parts of diagrams are very often well visible and it is clear how to insert the missing parts. It is especially true if a connector notation is used.

The diagrams are especially good during the initial steps of the solution of issues. Diagrams provide a powerful outline of a system provided it is not too large.

It is worth of mention that in all these cases the diagrams are used like figures or blueprints in technical and scientific publications and documents. They are in fact the part of the (text) document. The role of the diagrams is so important that the document is used during any update of the text the diagrams are updated too. The problem of obsoleting diagrams can be then avoided.

4. Diagrams in the Later Stages of Software Life Cycle

Diagrams can be used in the later software life cycle design through maintenance. Typical aims can be:

1. The enhancement (better quality) of user interfaces, i.e. the enhancement of system usability (compare [6]).

2. The better understanding of the system requirements by system designers, coders, testers, and, sometimes, maintainers. As diagrams are difficult to modify properly, they are not too useful for maintenance. It is true especially for complex diagrams and tasks.

3. Implementation of a tool to support decisions during design, coding, and sometimes testing.

4. Code generators. It is typical for model-driven architecture.

5. Auxiliary tools for design and coding and for code changes.

The use of diagrams in the ways described in 1 and 2 is rather a necessity than an option. Following applications of diagrams can have substantial positive effects:

- The use of diagrams as an auxiliary tool is reasonable and effective provided that the diagrams are discarded early.

- The practices mentioned in 3 are classical but not too satisfactory. Code is the most important document in classical practices. Changes are typically done in the code first and then, hopefully, in related diagrams. The changes in text are often easier than the changes in the diagrams. There is therefore no strong need to update the diagrams. The result is that the diagrams become obsolete and it ”does not matter” for some time. The effort needed to update the diagrams is then felt superfluous. The final state is that only code does matter – see the principles of agile development.

In large systems the diagrams are so complex that they lost the advantages discussed above. The use of complex diagrams can then become contraproductive.

5. The Case of Model Driven Architecture

The main issue with software development oriented diagrams is that they often are an auxiliary tool only. One solution is that diagrams with some additional information only are used and no open source code exists or it cannot be used.

The sample researches indicate that MDA are rarely used, compare [1] containing results of the research in Czech Republic. The reasons for such conclusions are:

1. The use of diagrams as a programming tool leads to decisions to use the diagrams as a specification mean. It leads to the antipattern ”premature programming” as the requirements are transformed into diagrams that need not be well suited to user knowledge domains and languages. The requirements are then usually not well formulated and yet worse transformed or adapted to fit MDA domain, not the user knowledge domain.

2. The underlying automated code generation system (ACG) must ultimatively be without errors. It is usually hopeless for developers to repair the failures of ACG or to change the generated code for other reasons, for example, effectiveness reasons. The errors in ACG superpose the errors in code compilers.

3. The (collection of the) diagrams necessary to model a given system is very complex. It is then quite difficult to navigate across the ”database” of diagrams. It can be e.g. quite uneasy to look for some names or patterns.

4. Some phenomena need not be easily described via MDA diagrams (e.g. some aspects of SOA service orchestration).

5. Small changes of the generated code can require large and laborious changes in the structure of diagrams.

6. The use of MDA requires painful changes in software development practices. On the other hand, it fixes current state of art for e.g. object-oriented attitude for a too long time.

7. Current MDA are on the other hand too object oriented. It can be disadvantageous if one wants to integrate batch systems or to design user-friendly interfaces, etc.

8. There is almost no guarantee that the life-time of the MDA supporting system will be long enough to enable a reliable support covering entire lifetime.

We can conclude that MDA is a promising concept but it is now well usable for smaller non-critical systems only.

6. Texts as well as Diagrams

Some stages of software development must use texts and diagrams. Examples are specification and architecture description. It is reasonable to attempt to find some advantages from it.

The most important advantage is the possibility to apply the general principles of the writing of well-formed documents. Such documents are in their text part reasonably structured and their ”graphical” part does not use cumbersome figures.

Plain text can be flexibly structured using standard methods like paragraphs, chapters, abbreviations, links, indexes, and so on. Changes can be made very easily and can be easily logged. It can be objected that text is not clear and illustrative enough. Note, however, that the clearness of the text need not deteriorate if the document size grows. This property is not observed for diagrams, the clearness of which falls with size.

Proper combination of texts and diagrams enables the more flexible structure of documents. It is a good practice for technical documents.

Texts can be now, using e.g. XML, structured in a very sophisticated way enabling e.g. very powerful document presentation in digital form. There is a lot of powerful tools for the text generation, looking for editing, etc. It is not too difficult to guard whether changes in the text were propagated into diagrams (pictures) and vice versa.

Such an attitude can substantially weaken the drawback of diagrams that there are no satisfactory tools enabling searching for diagrams having similar semantics.

All it works well for the diagrams not defining directly the program structure and used during system maintenance (like the diagrams used in MDA). In this case the only feasible solution seems to be full equivalence of diagrams and code. In other words the development of tools enabling the generation of the code from diagrams and vice versa the generation of the diagrams form the code provided that the code fulfills some standards. Such tools are not fully available, but some solutions exist. The transformation diagrams → code is available in MDA system. The transformation code → diagrams is known as reverse engineering. Available solutions are, however, not powerful enough. MDA diagrams tend to be too much programming oriented (see above).

Experiments with tools like ACASE [8] have shown that it is possible to make tools allowing to display code as text or as a diagram. Users let display the code as diagrams when working with simple algorithms. Complex algorithms were typically displayed as text as it is possible to see a larger part of the algorithm at once. Sometimes the combination has been used: The critical control structure has been shown as diagram, the rest as text.

It seems that it clearly demonstrates the usability of such tools: The beginners may start with more intuitive diagrams, complex things may be displayed as text, and finally, when it is necessary to analyze the code, the combination of code and diagram can give the highest benefit.

7. Conclusions

The use of diagrams and other graphical means during the development must be used as a tool supporting specification or a mean supporting the initial stages of a problem solving process. The problem of applications is the description of overall architecture and similar properties of systems.

In these cases the diagrams should be combined with text in a structured natural language. If made properly, the combination of texts and diagrams can bring great benefits as the advantages of both forms can be combined and their disadvantages eliminated. The existing CASE systems do not support enough such a solution.

The use of diagrams as in specification should be viewed as a use of diagrams as natural language ”enhancement”.

The application of diagrams to describe and maintain the structure of the system in the small, e.g. to define programming constructs so that they can be maintained has not been for large systems solved properly yet. It is not clear, whether the use of such detailed diagrams for such a purpose is even a reasonable goal.

Acknowledgement

This research was partially supported by the Program ”Information Society” under project 1ET100300517 and by the Czech Science Foundation by the grant number 201/09/0983.

References

[1] L. Bartoň. Properties of MDA and the ways of combination of MDA with other requirements specification techniques (in Czech). Master’s thesis, Charles University, Faculty of Mathematics and Physics, Prague, Czech Republic, 2006.

[2] K. Beck. *Extreme Programming Explained: Embrace Change*. Addison Wesley, Boston, 1999.

[3] K. Beck, M. Beedle, A. van Bennekum, A. Cockburn, W. Cunningham, M. Fowler, J. Grenning, J. Highsmith, A. Hunt, R. Jeffries, J. Kern, B. Marick, R. C. Martin, S. Mellor, K. Schwaber, J. Sutherland, and D. Thomas. Agile programming manifesto, 2001. http://www.agilemanifesto.org/.

[4] IDS Scheer. Aris process platform. http://www.ids-scheer.com/international/english/products/31207.

[5] J. Král. *Informační Systémy, (Information Systems, in Czech)*. Science, Veletiny, Czech Republic, 1998.

[6] J. Nielsen. *Usability Engineering*. Academic Press, New York, 1993.

[7] OMG. Model driven architecture, 2001. http://www.omg.org/mda.

[8] M. Žemlička, V. Brůha, M. Brunclík, P. Crha, J. Cuřín, S. Dědic, L. Marek, R. Ondruška, and D. Průša. Projekt Algoritmy (in Czech: The Algorithms project), 1998. Software project.

PSS: A Phonetic Search System for Short Text Documents

Jerry Jiaer Zhang Son T. Vuong
University of British Columbia, Canada
2366 Main Mall, Vancouver, B.C., Canada
(1)-604-822-6366
{jezhang,vuong}@cs.ubc.ca

Abstract

It is the aim of this paper to propose the design of a search system with phonetic matching for short text documents. It looks for documents in a document set based on not only the spellings but also their pronunciations. This is useful when a query contains spelling mistakes or a correctly spelled one does not return enough results. In such cases, phonetic matching can fix or tune up the original query by replacing some or all query words with the new ones that are phonetically similar, and hopefully achieve more hits. The system allows for single- and multiple-word queries to be matched to sound-like words or phrases contained in a document set and sort the results in terms of their relevance to the original queries. Our design differs from many existing systems in that, instead of relying heavily on a set of extensive prior user query logs, our system makes search decisions mostly based on a relatively small dictionary consisting of organized metadata. Therefore, given a set of new documents, the system can be deployed with them to provide the ability of phonetic search without having to accumulate enough historical user queries.

Index Term *– Phonetic Match, Search*

1. Introduction

With the ever increasing amount of data available on the Internet, quickly finding the right information is not always easy. Search engines are being continuously improved to better serve this goal. One useful and very popular feature is phonetic matching. Google's "Did you mean" detects spelling errors if not many matches are found and suggest the corrections that sound like the original keywords. Yahoo and MSN use the different names "We have included" and "Were you looking for", but they essentially do the very similar thing. This feature has become so popular that almost all the big search engines cannot run without it. However as much as it is highly demanded, not many websites can afford to provide this kind of user experience mostly due to the technical limitation: an extensive set of historical queries to build a statistic model for word retrieval and correction. PPS is to address this gap. It is a search system replying on a relatively small, self-contained dictionary with phonetic matching ability that is similar to what the big websites can offer.

In this paper, we propose the design of PPS, which only requires a small data set to function. It focuses on the correlations among different words and phrases, as well as the relationships between words and the containing documents to create a dictionary for phonetic searches on single- and multiple-, correctly spelled and misspelled words and phrases.

The remainder of the paper is organized as follows: Section 2 provides the system design. Section 3 takes a look on evaluation in terms of phonetic matching accuracy and efficiency. Section 4 concludes this paper and gives an outlook to future work.

2. System Design

This section presents two parts. The first is the creation of the dictionary data structure PPS relies on. The second is the phonetic matching mechanism based on the dictionary.

2.1. Dictionary Creation and Maintenance

We organize the data in a way that allows fast access, easy creation and maintenance. The data structure storing the documents serves as a dictionary for non-linear lookups. It also contains meta-data that describes document properties for multi-word sound-based searching.

2.1.1. Text Processing

Given a document, it is not difficult to break it into words. In PPS, we identify words with regular expressions to match predefined regex patterns to words. A set of unique words containing letters and digits are extracted from this process.

2.1.2. Dictionary Creation

The processed documents can then be used to create the dictionary that carries not only the original document text but also additional information that describes their properties. The following sections discuss the creation of these properties that are stored together with the original documents they are derived from as metadata.

Word List

A Word List is merely a list of distinct words that appear in the document retrieved during the text processing phase. It is sorted alphabetically. Sorting could be somewhat expensive but there are 2 reasons of doing it. First, documents tend to be static once they are stored in the database, so sorting usually only needs to be performed once for each document. Second, the overhead of dictionary creation does not add to the searching run time, so it is preferable to organize the data in a way that facilitates search performance. We can use binary search on the sorted list for word matching to achieve O(lgn) time complexity.

tf-idf Weight

tf-idf Weight is a statistical measure to evaluate the importance of a word to a document in a set of documents [1] [7]. It is obtained by multiplying Term frequency and Inverse Document Frequency. A high tf-idf weight is archived by a high term frequency in the given document and a low document term frequency in the whole set of documents. Therefore, terms appearing commonly in all documents or infrequently in a considered document tend to be given low weights and thus can be filtered out [8].

Double Metaphone Code

Double Metaphone indexes words by their pronunciations and generates two keys, primary and alternate, that represent the sound of the words [5]. To compare two words for a phonetic match, one takes the primary and alternate keys of the first word, compare them with those of the second word. The two words are considered phonetically matching only if their primary and/or alternate keys are equivalent [5].

Local Phrase Frequency

The local phrase frequency keeps track of the frequency of phrases in a document. To the context of this paper, a phrase is two or more consecutive words in the same order as they are in the contained document. We count phrase frequencies by grouping every two consecutive words and calculates the frequency and then grouping every three consecutive words and calculates the frequency. This process goes on till it groups all words of the document and calculates the frequency. Phrases derived from the above list are searched through the whole document to count their occurrences. To prevent bias towards longer documents, the occurrences are divided by the document's word length. The quotients thus serve as the phrase frequencies. Each phrase, together with its frequency, is then saved in a local phrase frequency table for each document. We call it local because this value is independent of the content of other documents in the document set.

Global Phrase Frequency

After the local phrase frequencies of a document are calculated, they are added to the global phrase frequency table. If the phrase exists in the table, its frequency is increased by the local phrase frequency. PPS uses it to determine how often a set of words occur together, as well as how frequently such a combination appears across documents.

2.1.3. Dictionary Maintenance

When new documents are added to the document set, the dictionary is updated to adjust the relative term match strength of each document that is derived from these documents. The major work is to re-calculate the tf-idf weight via a database script. It periodically processes the new documents since the last run and re-adjusts properties related to the whole document set at the end.

2.2. Single Word Search

Searching for a single word involves finding all matching documents and sorting them in the order of relevance. If the number of results is lower than the predefined configurable Result Size threshold, the system starts phonetic matching. Then results are ranked based on their relevance to the query and only those that exceed a predefined Sound-Like threshold are returned.

2.2.1. Word Matching

We use the Boolean Model to find matching documents Search is purely based on whether or not the query word exists in the document word lists. Boolean Model is quite efficient at this since it only needs to know whether or not the qualified documents contain the queried terms.

2.2.2. Result Sorting

The retrieved documents texts are represented as vectors in an algebraic model where each non-zero dimension corresponds to a distinct word in that document [6][7]. Building vectors for the respective documents can calculate the document similarities by comparing the angles between them [2]. If we compare the angles between a query and the retrieved documents, we can tell how "close" each document is to the query. A common approach to calculate vector angles is to take the union of the terms in two documents as the dimensions, each of which contains the frequency of the word in that document. PPS has improved it for better accuracy.

First, instead of using term frequency as values for vector dimensions, we applied the tf-idf weights to evaluate the importance of word to the considered document [7] because longer documents might have a low proportional term frequency even thought that term may have a higher occurrence than it does in a much shorter document. In such cases, it is imprudent to simply take the longer one. We apply tf-idf weights since the local tf parameter normalizes word frequencies in terms of the length of the document the words reside in. The global parameter idf contributes to the result the frequency of the documents containing the searching word relative to the whole document set. The product of the two parameters, the tf-idf weight thus represents the similarity of two documents with respect to the local term frequency ratio and the overall document frequency ratio [9]. In other words, rare terms have more weight than common terms. In our system a document is represented as a weight vector:

$$\mathrm{v} = [\mathrm{tf} - \mathrm{idf}_1, \mathrm{tf} - \mathrm{idf}_2, \mathrm{tf} - \mathrm{idf}_3, \dots, \mathrm{tf} - \mathrm{idf}_i,]$$

where i is the total number of distinct words in two documents.

Incorporating the above change, the sorting process works the following way:

1. Construct two initial document vectors of the same dimensions from the query and a document
2. Take the tf-idf weight values of the query and the document from the dictionary and fill them into the corresponding vector dimensions
3. Calculate the angel between the two vectors
4. Repeat step 1 to 3 for each document in the result set returned by Boolean text matching
5. Sort the result set by the cosine values of the angles. A larger number indicates higher relevance of the corresponding document

2.2.3. Phonetic Matching

If the above step does not return the documents the user looks for, PPS starts phonetic matching. The system first performs a match operation assuming the spelling is correct. If still not enough results are returned, then it performs another search operation with spelling correction.

Low Hits Resulted from a Correctly Spelled Query

PPS first tries to broaden the result by looking up sound-like words in the document set. Because words of same or similar pronunciations are encoded into the same or similar Double Metaphone code, a simple database query comparing the index Double Metaphone codes of two words will return a set of words that sound like the queried one. These words will be sorted by their pronunciations close to the original word by the Levenshtein Distance of their Double Metaphone codes. Because Double Metaphone codes are strings, we can apply the Levenshtein Distance to measure their differences and thus calculate the similarities of their sounds. Words that are phonetically identical always have the same Double Metaphone code, so their Levenshtein Distance is 0. As the pronunciations of two words become less and less alike, their Double Metaphone codes will have more different characters from each other and thus result in a further Levenshtein Distance. The system ranks the Levenshtein Distances between the query and the candidate words, and sorts them based on the different Levenshtein Distances.

Low Hits Resulted from an Incorrectly Spelled Query

If a query is misspelled, PPS first finds correctly spelled candidate words that are close to the query word, and then it ranks the candidates and returns the most matched one(s). The next two sections discuss each of the above steps in details.

Find Candidate Corrections: We observed that in most cases a misspelled word had a Levenshtein Distance of no more than 3 from the correct word. We also noticed that errors tend to occur towards the end of words. Because we are only interested in those that are close to the query word, the above two observations suggested that we could focus only on the 1-, 2-, and 3-Levenshtein Distances of the beginning portion of each word. The following is how it works:

1. Given a query word of length n, set $k = \lceil 0.6n \rceil$, where k is the size of leading characters to be taken from the query.
2. If $k \leq 3$, $k = \min(3, n)$; else if $k > 7, k = 7$. The lower bound of k guarantees there are enough permutations to form Levenshtein Distance of 3. The upper bound of k is 7. This reflects our observation that the beginning portion of a query word is more likely to be correct, so the following correction process will use this portion as the base for matching. The lower and upper bound of k were based on our experiments. They seemed to be the golden numbers that balanced accuracy and efficiency.
3. Take the first k characters of the query word and generate a key set where each item is a key whose Levenshtein Distance is 1, 2, or 3 from the k-length string.
4. Check each key in the key set for a given query against the word list metadata of each document in the document set. Return the words that also start with those same keys.

From our experiment, the size of candidate corrections only ranges from a couple of words to at most several hundred in a considerably big document set due to the large number of phonetically incorrect keys in the key set. Because of the relatively small data pool, we are able to implement a reasonably comprehensive scoring system to rank the candidates in order to find the best match.

Rank Candidate Corrections: Now that a list of candidate words have been found. The next step is to choose the best match(es). The system has a ranking system that takes into account the following factors:

Weighted Levenshtein Distance from a candidate to the original misspelled query word. The reason to compare it with the complete word rather than its first k characters is to ensure the evaluation reflects the relevance of a candidate to the query word as a whole. The concept has been commonly used in bioinformatics, known as the Needleman-Wunsch algorithm, for nucleotide sequence alignment [4]. It makes sense in our application domain because among all spelling mistakes, some are more likely to occur than the others. Table 2.1 is a list of considered operations and their costs in calculating the Weighted Levenshtein Distance.

Operation	Cost
Insertion	0.9
Deletion	1
Substitution	0.6
Transposition	0.6
Double Letter Error	0.4

Table 2.1: Operations for Weighted Levenshtein Distance calculation and their costs.

The cost associated with each operation was from our experiment. This combination seemed to produce better results than others. The Weighted Levenshtein Distance is the normalized total cost of performing these operations. If c is the total operation cost to transform a candidate to the query word, and n is the query word length, the score from the Weight Levenshtein Distance can be calculated as:

$$1 - \frac{c}{n}$$

where c is always less than or equal to n because the maximum cost is no greater than 1.

Next, *Starting and Ending Characters* of a candidate word are checked against those of the query word. The more beginning or ending characters the two words share in common, the more likely the candidate is the correction of the misspelled query. It was also from our tests the closer a letter is towards the middle of a word, the more likely a spelling mistake can happen. We took into account this factor in the ranking system with a linear scoring function which works the following way:

1. Set s = 0. From the first letter of the candidate and the query word, check if they are identical. If they are, increment s by 1 and move on to the next letter (in this case, second one) of both. Repeat this process until:
   a. the two letters at the same position from the two words are not the same,
   b. or the letter position is equal to half of the length of the shorter word.
2. Set e = 0. Starting from the last letter of the candidate and the query word, do the same as Step 1 except that it checks for the second half of the words.
3. The final score for this factor is calculated as:

$$\frac{s + e}{\min(n_q, n_c)}$$

where n_q is the length of the query word, and n_c is the length of the candidate word. The division is necessary to normalize the score to prevent bias toward longer words.

Third, the *Double Metaphone code* of both the candidate word and the query word are compared to calculate the third score based on their pronunciations:

1. if the primary key of the candidate is the same as the primary key of the query word, the candidate gets 0.3
2. else if the primary key of the candidate is the same as the alternate key of the query word, or if the alternate key of the candidate is the same as the primary key of the query word, the candidate gets 0.2
3. else if the alternate key of the candidate is the same as the alternate key of the query word, the candidate gets 0.1
4. else if none of the above three conditions is met, the candidate gets 0

The maximum score a candidate can possibly get from this factor is 0.3, which is lower than the other two factors. We made this decision based on two reasons. First, due to the complexity and the "maddeningly irrational spelling practices" [5] of English, the Double Metaphone algorithm may fail to generate unique codes to distinguish certain words. The second and more important reason is that, even if there were a perfect phonetic mapping algorithm that could distinguish every single different pronunciation, it is still not able to consider words that sound the same but differ in meanings. These words are known as homophone. Because it is unlikely that the users would misspell a word as one of its homophones, we had to be careful not to overly rely on phonetic similarity. This is why the Double Metaphone score is weighted only about 1/3 of the previous spelling-oriented factors.

Search on Phonetically Similar Word: Now that the best phonetically matched word is found, the system performs a single word search using the new word as the query word to find the result documents. This time, Phonetic Matching is performed because the new query is from the document set, which means a none-empty return set is guaranteed.

2.3 Multiple Word Search

Similar to single word search, there are two stages involved in multiple word search:

1. The system performs text matching search. If the queried phrase is found in more than the Result Size number of documents, the system sorts and returns all of them.
2. If the queries word is not found or only exists in the number of documents smaller than the Result Size threshold, the system performs phonetic matching search, sorts and returns the results.

There are also 3 steps in these stages like in single word search: Phrase Matching, Resulting Sorting, and Phonetic Matching but their implementations are somewhat different.

2.3.1. Phrase Matching

The Boolean Model is applied to check the phrase against the local phrase frequency tables. Since each table consists of entries of two or more words separated by a whitespace, the query also needs to be re-formatted to have exactly one whitespace in between words to make it more search-friendly. Furthermore, phrase entries whose word lengths are less than that of the query string are neglected because it is impossible for them to hold the query string.

2.3.2. Result Sorting

Sorting is based on the importance of the query string to both the document and the whole document set. Therefore, this is where both the local and the global phrase frequency tables are needed. Sorting for multi-word queries is actually much easier than for single-word queries because the Vector Space Model with tf-idf weights used for the single-word ones does not apply here due to the fact that counted frequencies of phrases are not all meaningful. For example, given the sentence "How are you", "how are" is a valid phrase in PPS but it does not function as a meaningful unit in the sentence. On the other hand, from our tests, the simple phrase frequency comparison worked well. Each document gets a score which is the product of the local and global phrase frequencies of the query string. The higher the score is, the more relevant that document is to the query string. This method produces reasonably good results because it takes into account the importance of a phrase both locally to the document and globally to the whole document set.

2.3.3. Phonetic Matching

Similar to single-word search, if the strict text-based matching does not return satisfying results for the phrase, the system starts the sound-based search following these steps:

1. Break a query phrase into a list of single words.
2. For each word, perform the single-word phonetic matching operation to retrieve a list of top candidates.
3. Consider all possible permutations of the candidate lists by taking one word from each of them. For each permutation, refer to the global phrase frequency table to get its global frequency in the whole document set. This is called correlation check.
4. After all permutations are generated and their global phrase frequencies are check, return the one with the highest frequency.

As the query size increases, the permutations from all candidate word lists grow exponentially. Fortunately, we observed that a permutation could be generated by selecting one word from each candidate list and then concatenating the selections together. It means before a permutation is formed, all entries in the global phrase frequency table are possible matches. Then, the first word from the first candidate list is chosen as the first element of the permutation. At this point, those phrase frequency entries not containing the same first word can be purged. Next, the second word from the second candidate list is chosen as the second element of the permutation. Among the phrase frequency entries left from the previous selection, those without the same second word can also be purged because there will be no match to the whole permutation for sure. The process goes on till either the permutation is completed or there is no phrase frequency entry left. If the permutation is completed, it means there is a match in the phrase frequency table. Otherwise, there is no such a phrase that can match the incomplete permutation from its first element up to its last that is generated right before the process stops. Therefore, all permutations with the same beginning elements as the incomplete one can also be purged. Moreover, we can further optimize the process by reducing the phrase frequency pool to only those entries with the same number of words as that of a complete permutation. This limits the initial data set size to make it converge more quickly. Figure 1 is an example of the optimized permutation generation process on a three-word query.

Suppose there are five three-word phrases in the phrase frequency table. They are “A C A”, “B A B”, “B C A”, “C C A”, and “C C B”. Regardless of the size of the original phrase frequency table, these five are always the ones to beginning with because any other phrases with more or less words are purged. For simplicity, each candidate list has four words, “A”, “B”, “C” and “D”, to be chosen from, and there are three such candidate lists. Therefore, without optimization, a total of 64 (4 · 4 · 4) permutations are needed for the five existing phrases.

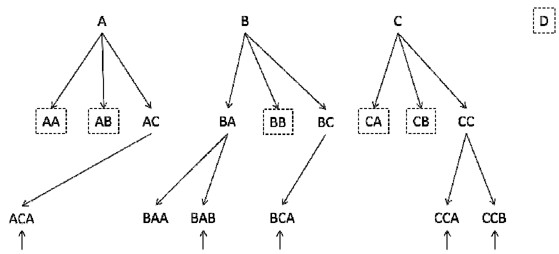


Figure 2.1: Optimized permutation generation process on a three-word query

Let’s see how optimization can speed up this process. We begin by picking “A”, the first word in the first list, as the first element of the permutation. Because there is only one phrase of the five starting with “A”, the other four do not need to be checked for the rest of this permutation. Then another “A”, the first word in the second list, is picked as the second element of the permutation. Now this permutation starts with the words “A A”. Because the only phrase left from the last selection does not start with “A A”, we can stop this permutation and any other permutations starting with “A A”. In Figure 1, “A A” is surrounded by a dotted board to represent the termination of this “branch”. Next, we pick the second word “B” from the second list. Similarly, there are no phrases start with “A B”, so any permutations of “A B X”, where “X” can be either “A”, “B”, “C” or “D”, are ignored. Thus, “A B” is also surrounded by a dotted board. Next, we pick “C” from the second list to form “A C”. Because “A C A” matches to “A C” for now, we can move on to the third list and select “A” from it to form the first complete permutation “A C A”. At this point, we find a match and no further permutations of “A C X” will be performed because we know there is only one phrase in the form of “A C X”. Instead of 16 permutations and comparisons, only one is generated and 5 comparisons are made between the incomplete permutation and the phrase entries for all permutations starting with “A”. Repeat the same steps for the rest till all phrases are found. One extreme case is when a permutation starts with “D”. All the 16 “D X X” combinations are ignored. A save of 16 generations and 15 comparisons! In figure one, only a total of 6 complete permutations are generated and 5 of them are the matches. From out tests, such optimization could save over 90% of time on average.

3. Evaluation

This section discusses the performance of PPS. A simulator and test data were created to search for restaurant names throughout the Greater Vancouver Region. We examine the effect of single- and multiple-word searches with phonetic matching. By comparing the results to the actual data in the test document set, we evaluate the search accuracy and running time of the system with different types of inputs.

3.1 Simulation Setup

3.1.1 Test Data Pool

The test data is a restaurant names in the Great Vancouver Region. We have built a crawler with the free software Web Scraper Lite [3] to grab and extract restaurant listings from Yellowpages.ca into MySQL. The data pool consists of more than 3800 restaurant names. We chose restaurant names as our test data because of two reasons. First, PPS was designed specifically for short text documents. The lengths of restaurant names usually varied from one to eight words, and thus would make good test data for the evaluation. Secondly, a lot of the names were non-English so phonetic matching would be useful.

3.1.2 Test Input

We created the test input in two stages. First, a set of correctly spelled words and phrases were generated. These words and phrases must not appear in the test data pool. Second, we created a set of misspelled words and phrases with a Levenshtein Distance greater than zero but less than or equal to five from the existing words and phrases in the test data pool. There are 1000 inputs in total for the test. Table 3.1 is a summary of the types and the sizes of the input we tested on.

Input Type	LD1	LD2	LD3	LD4	LD5	Total
Correct Word	N/A	N/A	N/A	N/A	N/A	250
Correct Phrase	N/A	N/A	N/A	N/A	N/A	250
Misspelled Word	61	59	49	44	38	250
Misspelled Word	59	55	53	46	37	250

Table 3.1: Test input types and sizes

3.1.3 Simulator

We implemented a simulator in PHP to query the test data pool with the test inputs and to collect the test results. What it essentially does is the two searching stages for single- and multiple-word queries described in the previous chapter.

3.2 Simulation Results

The primary goal of the simulation is to evaluate the accuracy of phonetic search when dealing with different types of input: correct word, correct phrase, misspelled word, and misspelled phrase. We will discuss each of them in this section.

Input Type	# of Queries	# of Matches	%
Correct Word	250	239	95.6
Correct Phrase	250	216	86.4
Misspelled Word	250	223	89.2
Misspelled Phrase	250	212	84.8

Table 3.3: Number of phonetic matches from the correct word, phrase queries, and from misspelled word, phrase queries

The first two rows of Table 3.3 are the search results when a query was a correctly spelled word or phrase. The system yielded a 95.6% successful rate when dealing with single-word queries. It was because the search process took a regressing pattern to gradually increase the Levenshtein Distance between the Double Metaphone code of the query word and that of a document until it found the first match. For those the system did not find a match, it was because they were generated so randomly that their Double Metaphone Levenshtein Distance from any document was equal to the length of the Double Metaphone code itself. In other words, these words did not sound like any words in the test data. Searching for correctly spelled phrases yielded a lower 86.4% successful rate. This is because the system needs to find candidate words that are phonetically close to every word in a query phrase. If any word returns an empty candidate list, the matching stops. Furthermore, the more words a query phrase has, the less likely there is a match in the document set. This observation was proven by the fact that, among the 13.6% unsuccessful query phrases, most of them consisted of five or more words.

The last two rows of Table 3.3 are the search results when a query word or phrase was misspelled. Single word queries yielded a high 89.2% successful rate. When we were generating the test input, we intentionally made all queries a Levenshtein Distance no more than 5 to model the common error patterns. This is why the phonetic matching worked well with spelling mistakes. It came a little surprise that the unsuccessful words were the smaller ones. We think it was because after normalization, even close Levenshtein Distance could be proportionally large to small words. For misspelled phrases, the successful rate is close to its correctly spelled counterpart. This was expected because the decisions were made on the same factors - the Local and Global Frequencies.

Input Type	Queries	Min.	Max.	Avg.	Std. Dev.
Text-based Word Search	100/100	110	371	193	76
Text-based Phrase Search	100/100	216	478	305	93
Sound-based Correct Word Search	239/250	607	1131	866	137
Sound-based Correct Phrase Search	216/250	715	1206	893	99
Sound-based Mis-spelled Word Search	223/250	975	1704	1252	281
Sound-based Mis-spelled Phrase Search	212/250	874	1569	1174	325

Table 3.4: Running time of different types of searches

Table3.4 shows the running time of all six different types of searches. It was no surprise that the simplest text-based searches took the least time while the sound-based misspelled word and phrase searches took the longest. It is worth mentioning that the maximum average search time is merely over a second and all types of searches have a small standard deviation comparing to its average time. Combining Table 3.3, we conclude that the system has behaved reasonably fast and stable with an over 80% successful rate regardless of the various types of inputs.

4. Conclusion and Future Work

In this paper, we introduce PPS, a search system based on both text and sound matching for short text documents. The system makes incorporates several widely adapted algorithms into its staged searching process to deal with different search cases. Each stage has its own scoring model built upon some common algorithms and the metadata specifically prepared for it. The various metadata associated with documents are the keys to the dictionary-based approach our system takes for phonetic searching. We provide a high level design specifying the system implementation from dictionary creation and maintenance to text- and sound-based matching for various types of queries. We also evaluate the system performance under these circumstances. The results suggest that our system meets its design goal with respect to accuracy and efficiency.

There are several areas in the development of the system that deserve further exploration. First of all, stopwords like “the”, “to”, “not” or “is” appear much more often than others but carry very little information. Building dictionary metadata for them is expensive and usually useless. It could be helpful to skip these stopwords without sacrificing correctness in matching a phrase like “to be or not to be” that contains only stopwords. Secondly, when searching for misspelled words, the current design does not take missing whitespaces into account. Consider the word “georgebush”. PPS would return something like “Georgetown” while a better match might be “George Bush”, which would be found by inserting whitespaces into the keyword. Similarly, we can also consider combining words. Again, the challenge here is to find the right granularity to balance between accuracy and efficiency.

References

[1] William B. Frakes and Ricardo A. Baeza-Yates, editors. Information Retrieval: Data Structures & Algorithms. Prentice-Hall, 1992.

[2] Yuhua Li, Zuhair A. Bandar, and David McLean. An approach for measuring semantic similarity between words using multiple information sources. IEEE Transactions on Knowledge and Data Engineering, 15(4):871-882, 2003.

[3] Web Scraper Lite. http://www.velocityscape.com/products/webscraperlite.aspx

[4] Saul B. Needleman and Christian D. Wunsch. A general method applicable to the search for similarities in the amino-acid sequence of two proteins. Journal of Molecular Biology, 48(3):443-453, March 1970.

[5] Lawrence Philips. The double metaphone search algorithm. C/C++ Users J., 18(6):38-43, 2000.

[6] Vijay V. Raghavan and S. K. M. Wong. A critical analysis of vector space mode for information retrieval (5): 279-287, 1986.

[7] G. Salton, A. Wong, and C. S. Yang. A vector space model for automatic indexing. Commun. ACM, 18 (11):613 -620, 1975.

[8] Gerard Salton and Christopher Buckley. Term-weighting approaches in automatic text retrieval. In Information Processing and Management, pages 513-523, 1988.

[9] Gerard Salton and Michael J. McGill. Introduction to Modern Information Retrieval. McGraw-Hill, Inc., New York, NY, USA, 1986

Hybrid Client-Server Multimedia Streaming Assisted by Unreliable Peers

Samuel L. V. Mello, Elias P. Duarte Jr.
Dept. Informatics – Federal University of Parana – Curitiba, Brazil
E-mail: {slucas,elias}@inf.ufpr.br

Abstract

Stream distribution is one of the key applications of the current Internet. In the traditional client-server model the amount of bandwidth required at the streaming source can quickly become a performance bottleneck as the number of users increases. Using peer-to-peer networks for distributing streams avoids the traffic concentration but, on the other hand, poses new challenges as peers can be unreliable, presenting a highly dynamic behavior, leaving the system at any time without previous notice. This work presents a hybrid strategy that uses the set of clients as an unreliable P2P network to assist the distribution of streaming data. A client can always return to the server whenever peers do not work as expected. A system prototype was implemented and experimental results show a significant reduction of the network traffic at the content source. Experiments also show the behavior of the system in the presence of peer crashes.

1. Introduction

The efficient and reliable distribution of live multimedia streams to a very large number of users conneted to a wide area network, such as a corporate WAN or the Internet itself, is still an open problem. As the traditional client-server paradigm does not offer the required scalability, peer-to-peer (P2P) networks have been increasingly investigated as an alternative for this type of content distribution [1, 3, 6, 7].

In the client-server model, the server is the central data source from which data is obtained by clients. All transmissions necessarily demand that clients connect to the server, which can easily become a performance bottleneck. On the other hand, the client-server model also has advantages, such as the ability of the content owner to control the delivery, for instance blocking some clients or employing a policy-based delivery strategy. This property is especially welcome when different clients have different quality of service requirements. Several on-line providers currently deliver live audio and video streams using tools based on the client-server paradigm.

In this work we propose a client-server multimedia streaming strategy that employs a P2P network for reducing the bandwidth requirements at the server. Typically, a server transmit the data to the several clients in parallel. As clients receive the content, they build a peer-to-peer stream distribution network and exchange different pieces of the received data. This approach is similar to the one employed by BitTorrent [4]. The server delivers different parts of the data to the clients, each of which is eventually able to obtain the whole content by exchanging its received data with other clients that received different parts of the same content.

In live multimedia streaming systems the content is produced at the server in real time and must be available to the clients within a maximum time limit, in order to be successfully used by an application that consumes the data at a fixed rate. In this case "old" data is not useful for clients. This implies that the variety of parts that are delivered to clients that start receiving the content at roughly the same time is small, as each client typically keeps data spanning at most a few minutes in its local buffer.

The proposed multimedia content distribution system is actually hybrid, as it presents characteristics both of the client-server and P2P paradigms. The server is responsible for creating the content, temporarily storing it in a local buffer, splitting the buffer in small parts and transmitting these parts to a group of clients. These clients interact among themselves and with other clients that did not receive the data directly from the server to set up forwarding agreements to exchange the received data. The forwarding agreements are refreshed from time to time to adjust to changes in the system. We assume that the the P2P system is dynamic and has very low reliability, as clients acting as peers can leave the system at any time. As the client buffer must be ready for playback within a bounded time interval, whenever a peer from which a client was supposed to obtain content fails or leaves the system, the client can obtain that piece of content directly from the server and within a large enough time frame so that data playback is not compromised. The proposed approach was implemented as a prototype, and experimental results are described showing the improvement on the bandwidth requirements at the server, as well as the system robustness in the presence of peer failures and departures.

Related work includes several P2P multimedia streaming strategies, such as [2, 3] that are modified versions of the BitTorrent protocol for continuous stream distribution; [5] also presents a system that is also similar to BitTorrent but employs network coding on the stream. Another common approach is to rely on a multicast tree for delivering the stream, such as in [6]. Some strategies such as [1] focus on VoD (Video on Demand), in which the user can execute functions on the stream such as pausing or fast forwarding. Usually pure P2P strategies do not offer QoS guarantees, such as [7] in which peers are selected based on their measured availability. The proposed approach is different from pure P2P strategies because the server remains responsible for eventually sending the stream if it is not received by clients from peers. Furthermore the server still interacts with all clients at least once every at every round in which peer agreements are established.

The rest of this paper paper is organized as follows. In section 2 the proposed strategy is described. Section 3 presents the implementation and experimental results. Section 4 concludes the paper.

2. The Proposed Hybrid Streaming Strategy

The proposed system has the basic components of the traditional client-server model. The server generates and sends the stream to the clients. The clients receive the stream and play it back to user. They also act as peers exchanging parts of the stream. The server helps clients to find one another to exchange parts of the data. In this work the terms client, node and peer are used interchangeably – but "peer" is more often employed for a client that is using its upload facilities to send part of the stream to another client. Each client keeps a list of peers from which it tries try to retrieve parts of the stream. If, by any reason, a client is not able to retrieve a given part from any peer, it returns to the server which provides the missing part. The stream is divided in *slices* that have fixed size and are sequentially identified. Slices are produced and consumed at a fixed rate. A slice is further divided in *blocks* which have constant size. A block is identified by the slice identifier and its offset within the slice. The stream is actually transmitted in blocks.

Clients send blocks to their peers according to forwarding agreements they establish. Each forwarding agreement specifies the transmission of blocks with the same identifier, for a certain number of slices. All forwarding agreements last the same number of slices, and each agreement starts at slices whose identifier is multiple of this number. Agreements are established in rounds, a round starts after all the slices of the previous agreement have been completely transmitted. The first slice of each agreement is called its *base-slice*. Whenever a client is unable to establish agreements with other peers, it still can establish an agreement with the server itself. Figure 1 depicts the blocks to be transferred in an example forwarding agreement. In the example, the current node receives from Node X blocks with identifier 3, starting from slice 20 and an agreement lasts 10 slices.

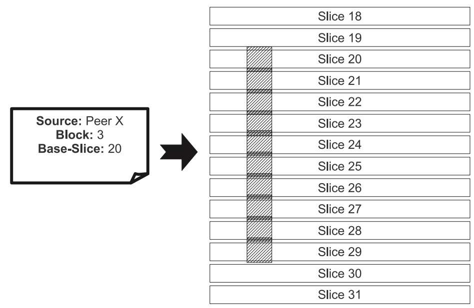


Figure 1. Forwarding agreement example.

Figure 2 shows the transmission of a slice. At the server, the media producer generates the content which is then divided in slices (1). Each slice has a sequential number and is further divided in blocks (2). In the example, each slice is divided in four blocks. The blocks are then transmitted to the clients (3). This transmission may take place directly from the server to the client or from a peer. Each block may take a different route to the client. After receiving the blocks, the client rebuilds the slice (4) and plays it back to the user (5).

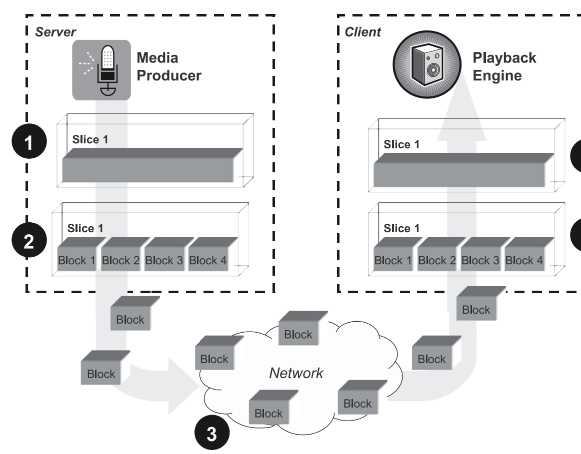


Figure 2. Blocks form the basic transmission unit.

The rate at which the slices are generated and their size are configurable parameters of the system. After a slice is sent to the clients, the server still keeps it stored in a local buffer, where it is available in case the server receives a retransmission request. The slices are held in this buffer at least until the next agreement is established. Figure 3 shows an example message flow as a client starts up. To make it simple, each slice is divided in only two blocks. When the client initializes, it registers itself at the server (1). The server accepts the connection and adds the client to a list of active clients. After that, the server sends back to the client information about the stream (2). This information is used to configure the client's playback engine and may include, for instance, the rate a slice should be consumed at and the size of each block. Afterwards, the client sends a request for information about other peers and the server replies with a set of peer identifiers randomly chosen from the active client list (3). Besides helping clients to find each other, the server also sends blocks to directly to several clients. In particular, after a new client starts up, the server creates forwarding agreements for all parts of the slice. The agreements are valid until in the next round of agreements the client can try to establish agreements with other peers (4); in this way a new client quickly starts receiving the stream.

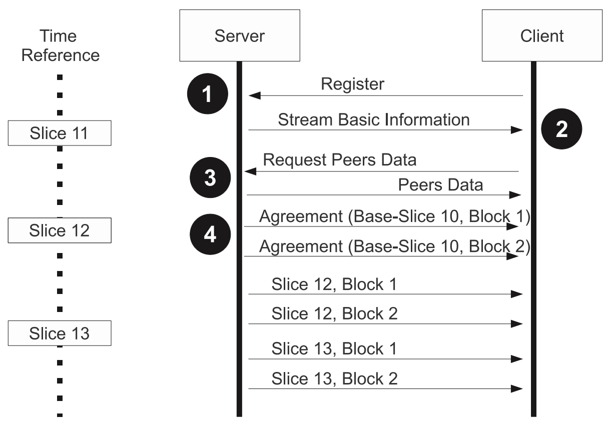


Figure 3. Message flow as a client starts up.

As the stream is generated at the server, it also has to send all the blocks to a subset of the peers, which then exchange these blocks so that all clients will eventually receive the complete stream. The largest the number of blocks sent by the server directly to the clients, the easier it is for each peer to find other peers from which it can receive the blocks it needs. We say that the server *spontaneously* chooses a number of clients with which direct agreements are established. This number of clients is usually a fraction of the total number of connected clients, and is a configurable parameter. The server can employ different approaches for selecting the clients for those agreements, for instance the selection can be based on the on RTT (Round Trip Time) to the client.

The clients that establish forwarding agreements directly with the server for a given base-slice are selected before a new round of agreements start. In this way, peers can establish the new agreements before the next base-slice gets transmitted. After choosing the peer that will receive the spontaneous agreements, the server notifies these peers with information about the blocks they are going to receive. The server then notifies all connected clients that there are forwarding agreements available for the next round and the peers can begin the agreement establishment phase as described below.

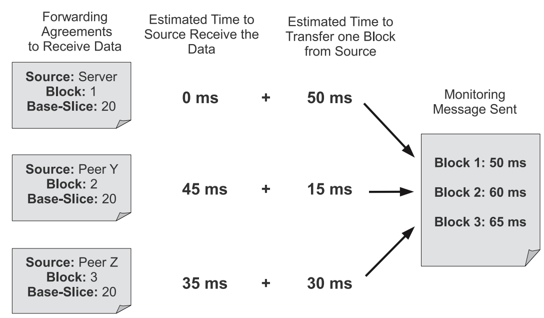


Figure 4. Monitoring message.

Every client keeps for each block a list of peers from which the client can try to establish an agreement to receive the block. This information is also used by a monitoring procedure. Periodically, each peer exchanges monitoring messages with the other peers in their lists. These messages contain approximate estimations of the delay between the creation of a new slice at the server and the expected arrival at the peer. As each block may go through a different route, there are different estimations for each block, as shown in the example in figure 4. Monitoring messages are padded so that they have exactly the same size of one block of the stream, so the peer can measure the time spent to retrieve a block (monitoring message) from each peer. This measures are taken at the agreement establishment phase.

When a client receives a notification from the server that there are forwarding agreements available for the next round, it creates a new peer list for each block. Each list is sorted by the estimated time to receive the data from the peer, computed as the sum of the time informed by the peer in its monitoring message and the time spent to retrieve the monitoring message from the peer. Furthermore, the client starts a timer that shows the end of the agreement establishment phase.

2.1. Agreement Establishment Phase

The client begins the agreement establishment phase sending a request to the first peer in each block peer list. If the peer accepts the request, the agreement is complete for that block until the next round of agreements is started. If the peer rejects the request, the client sends a request to the next neighbor in the list. The peers that reject the request are moved to the tail of the list, so another request is sent again if no other peer replies positively. As a node cannot accept an agreement to forward a block for which it does not have itself an established agreement, the delay between two requests may be enough for the peer to have the agreement and thus be able to to accept the request. This process is repeated until the client obtains forwarding agreements for all blocks or a timer expires showing that the agreement phase is over. If at the end of the agreement phase the client was unable to establish forwarding agreements for a block, it sends a request directly to the server. In this case, the client also sends a request for information about more peers to expand its block lists, in order to have a larger number of peers to try to establish agreements with in the next round, increasing the chance of success.

When a client receives an agreement request, it checks whether it has already established an agreement for the specified block and base-slice with another peer or the server itself. If there is such agreement, the client checks the number of peers with which it has already accepted agreements to forward the block to. Each client accepts at least a maximum number of forwarding agreements. If this number has not been reached yet, then the request is accepted, otherwise it is rejected. This maximum number to agreements that a peer can accept is a configurable parameter of the algorithm. Figure 5 shows an example of the agreement establishment phase wich is executed by every client after it receives the notification from the server. In the example, node A received an spontaneous agreement from the server for block 1, node Y for block 3 and node Z for block 2. Afterwards, node A creates (1) a peer list for block 2 and another peer list for block 3. Node A sends an initial request for block 2 to peer Y (2), which is unable to accept the request as it does not have an agreement to receive that data. The request is then sent to node Z, the next in the list (3). Node Z accepts the agreement and node A adds the agreement (4) to its list of established agreements.

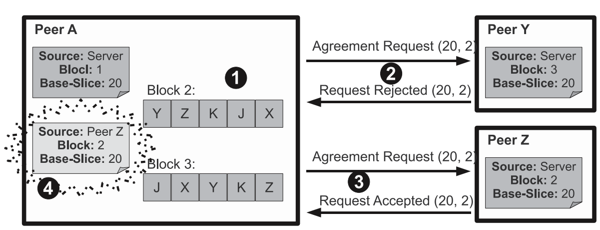


Figure 5. Agreement establishment.

2.2. System Behavior in the Presence of Peer Crashes or Departures

The slices received by the clients are stored in a buffer, from where they are consumed by the playback engine at a constant rate after an inicial delay. This initial delay gives a certain flexibility on the arrival times of different blocks. Before consuming a slice, the client must ensure that all blocks have been correctly received. If a block is missing, for example because the peer the block was supposed to come from has crashed or left the system, the client requests the missing block directly to the server. This process must be performed early enough so that all missing blocks can be retrieved from the server.

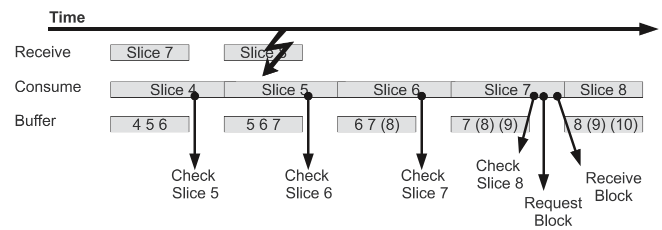


Figure 6. System behavior in the presence of peer crashes.

Figure 6 shows an example. A peer crashes and does not send blocks with a certain offset after slice 7. Slices from 8 to 10 are thus incomplete. Before consuming each slice, the client requests the missing blocks to the server, and only consumes the slice after the missing blocks are received. This procedure is repeated for all slices until the next round, in which the client will establish new forwarding agreements.

3. Experimental Results

This section describes an implementation of the proposed hybrid system and experimental results. The system was implemented in Java, with all messages exchanged by peers modeled as Java objects that are serialized and transmitted over TCP/IP connections. All network operations are handled by a wrapper class that also allows artificial bandwidth limits to be set and collects statistics on the amount of data sent and received. These artificial bandwidth limits allow the simulation of several client instances running at same host. All experiments involved the transmission of a 128 Kbps stream. A slice is composed of 8 blocks with 4 KBytes each, these blocks are played back each 2 seconds. Forwarding agreements last for 10 slices and the server makes spontaneous forwarding agreements with 40% of the clients after sending the fifth slice of each agreement. The

# Agrmts	8 Peers	16 Peers	32 Peers
1	5637 KB	8049 KB	12809 KB
2	5620 KB	10803 KB	15249 KB
3	6125 KB	13486 KB	17982 KB
4	6380 KB	16186 KB	20710 KB
5	6880 KB	18853 KB	23297 KB
Pure Client-Server	23040 KB	46080 KB	92160 KB

Table 1. Server bandwidth requirement given the number of spontaneous agreements.

number of times the same data was transmitted by the server was varied in the experiments. All clients were configured to have an artificial bandwidth limit of 1024 Kbps. The experiments were run several times and the results shown are representative of the values obtained.

The rest of this section describes three experiments, in which the following metrics were evaluated: (1) the reduction of server bandwidth required, (2) the influence of the number of copies sent by the sever on the upload bandwidth requirement at the clients, and (3) the system behavior in the presence of peer crashes.

3.1. The Reduction of Server Bandwidth Required

The first experiment shows the bandwidth requirement at the server as the number of spontaneous agreements with clients vary. From 1 to 5 agreements were established with 8, 16 and 32 clients. The system was executed in each case for 180 seconds. The results were measured at the server and consider all data sent, including control messages. In all cases, the playback interruption rate is below 1%, that is, more than 99% of the slices were available for playback at the expected time. Results are shown in figure 7 and table 1. For the varying number of clients, the total amount of data sent by the server is shown. The last column shows the minimum amount of data that would be transmitted using a pure client-server approach.

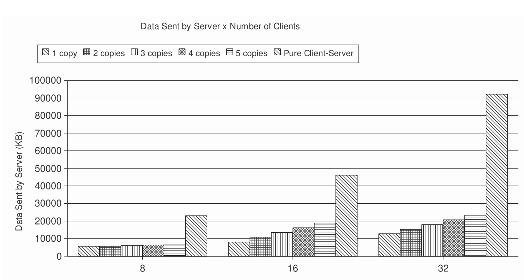


Figure 7. Server bandwidth requirement given the number of spontaneous agreements.

Copies	8 Peers	16 Peers	32 Peers
1	3313 KB	3639 KB	3943 KB
2	3302 KB	3523 KB	3805 KB
3	3297 KB	3366 KB	3692 KB
4	3317 KB	3201 KB	3657 KB
5	3292 KB	3029 KB	3620 KB

Table 2. Number of bytes sent by peers given the server spontaneous agreements.

3.2. Peer Upload Bandwidth Given Versus Spontaneous Server Agreements

The largest the number of spontaneous agreements the server establishes with the clients, the less data the peers need to upload. The graphic in figure 8 and table 2 show the average amount of data peers upload as the number of spontaneous agreements established by the server vary. Figures are an average of the amount data sent by each peer and include control and monitoring messages. It is possible to note that for 16 and 32 clients there is a slight reduction in the upload requirements. For 8 clients the value remained nearly constant, as the copies sent by the server are distributed among only 40% of the clients. Using these parameters, it is possible to see that a system with only 8 clients does not take advantage of the additional copies sent by the server.

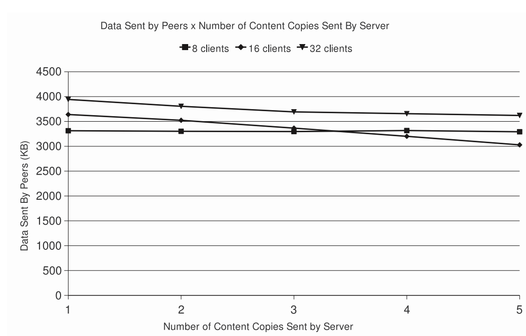


Figure 8. Number of bytes sent by peers given the server spontaneous agreements.

3.3. System Behavior in the Presence of Peer Crashes

The third experiment shows the system behavior in the presence of peer crashes. The system had the same configuration of the previous experiment, and the server established 3 spontaneous forwarding agreements. The network was composed of 32 peers of which 16 randomly crashed at a different instants of time. Clients who were supposed to receive data from a peer that crashed used the retransmission mechanism to request the missing blocks to the server. The system was able to keep the playback interruption rate below 1% at all working clients, even in the presence of failures of half of all peers, in other words, more than 99% of the slices were available when needed. The retransmission of missing blocks increased the bandwidth usage at the server, as shown in figure 9 and discussed below.

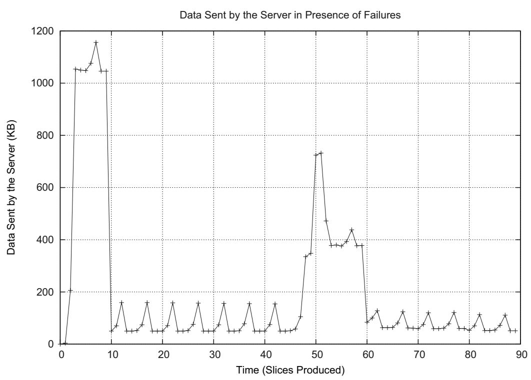


Figure 9. Required bandwidth at the server in the presence of peer crashes.

In figure 9 it is possible to note that a large amount of data is transmitted from the beginning up to the creation of slice 10. This high volume reflects the initial data sent by the server to new clients after the connection is established. In the experiment, all clients started at the same time in the very beginning of the run. Nevertheless, in real world scenarios, the clients are expected to connect at different time instants, and this reduces the amount of data the server has to send to all clients simultaneously.

Clients begin to take advantage of the forwarding agreements after the creation of slice 10. Then, the system presents a stable behavior up to the point a peer crashes, which occurs close to slice 45. During this time, it is possible to note some small peaks caused by monitoring and control messages. After the period where the system was under the effect of the failure, it is possible to note that those peaks are smaller, as the number of active clients exchanging control and monitoring messages with the server has reduced.

Although the crash occurs close to slice 45, but the effects on the server are observed only close to slice 48. The reason for this delay is that clients were still playing back previously buffered data. When the buffer gets nearly empty, the clients request missing blocks to the server. The experiment continues and peers keep on crashing until slice 60, when a new round of agreements will take place.

4. Conclusion

This paper presented a hybrid client-server multimedia streaming system which is assisted by an unreliable P2P network formed by clients. Clients receiving the content establish forwarding agreements for parts of the stream, avoiding the concentration of network traffic at the source. Clients can always return to the server after they are unable to retrieve the stream from a peer. Experiments show that the system provides an expressive reduction in the bandwidth requirements at the server and that the system is at the same time able to support peer crashes and departures without a significant interruption of the playback at working clients.

Future work include allowing clients to have different QoS requirements and also different bandwidth limits for downloading and uploading data. In the proposed strategy, clients are randomly selected as peers, this can be improved for instance by using a location aware peer selection strategy. The prototype we implemented allowed basic experiments to be performed showing that the strategy does work as expected, nevertheless large scale experiments must still be run comparing the system with other pure P2P streaming systems, with also check the limits on the system scalability.

References

[1] Y. Huang, T. Fu, D. Chiu, J. C. Lui, C. Huang, "Challenges, Design and Analysis of a Large-Scale P2P-VoD System," *SIGCOMM Comput. Commun. Rev.*, Vol. 38, No. 4, 2008.

[2] N. Carlsson, D. L. Eager, "Peer-assisted On-demand Streaming of Stored Media Using BitTorrent-like Protocols," *IFIP/TC6 Networking*, 2007.

[3] P. Shah, and J.-F. Paris, "Peer-to-Peer Multimedia Streaming Using BitTorrent," *IEEE Int. Performance, Computing and Communications Conference*, 2007.

[4] B. Cohen, "Incentives Build Robustness in BitTorrent," *Workshop on Economics of Peer-to-Peer Systems*, 2003.

[5] C. Gkantsidis and P. Rodriguez, "Network Coding for Large Scale Content Distribution," *INFOCOM*, 2005.

[6] D. A. Tran, K. A. Hua, T. Do, "ZIGZAG: An Efficient Peer-to-Peer Scheme for Media Streaming," *22nd Annual Joint Conf. of the IEEE Comp. and Comm. Societies*, 2003.

[7] X. Zhang, J. Liu, B. Li, T. P. Yum. "CoolStreaming/DONet: A Data-driven Overlay Network for Peer-to-Peer Live Media Streaming," *INFOCOM*, 2005.

Visual Programming of Content Processing Grid

Pierfrancesco Bellini, Ivan Bruno, Paolo Nesi
DISIT-DSI, Distributed Systems and Internet Technology Lab
Dipartimento di Sistemi e Informatica, Università degli Studi di Firenze, Firenze, Italy
http://www.disit.dsi.unifi.it/ , nesi@dsi.unifi.it, paolo.nesi@unifi.it

Abstract

The programming of GRID for content processing is a quite complex activity since it involves capabilities of semantics processing and reasoning about knowledge and descriptors of content, users, advertising, devices, communities, etc., and functional/coding data processing in an efficient manner on a Grid scalable structure. In this paper, the formal model and tool to visually specify rule programming on grid is presented. The tool has been developed on the basis of the AXMEDIS framework and grid tool, while it can be extended to support other formalisms generating processes for other grid solutions.

1. Introduction

With the introduction of User Generated Content, UGC, the back offices for content processing based on grid solution fulfilled the need to be more intelligent, flexible and scalable to satisfy quickly growing applications such as the back office activities of social networks. The grid computing provides high performance applications and resources widely distributed. These functionalities are becoming mandatory for web portals and end-users' applications. End–users' Grid are frequently used by non skilled users with no professional background on computer programming. For them, building or modifying grid applications is a difficult and time-consuming task. To build new applications, end users need to deal with excessive details of low-level APIs that are often platform-specific and too complex for them [1].

Some programming strategies and methodologies were proposed to encourage grid for end-users. The Problem Solving Environment (PSE) or portal [2], [3], makes the use of the Grid easier by supplying a repository of ready-to-use applications that can be reused by defining different inputs. Grid complexities are hidden, thus allowing only simple tasks (e.g., job submissions, status job checking). This solution does not provide the required flexibility to create real applications such as those needed to compose different services and processes.

Alternatively, the construction of scalable applications could be done by using workflow of services [4], [5], [6], [7]. These solutions are based on Workflow Management Systems, WfMS, and languages due to the relationships between the Grid solution and the Workflow solutions. On such grounds, they are unsuitable for semantic processing. Therefore, a tool to define visually the activity flow while combining basic processes and integration aspects of communication, flow processing, communication, data processing and semantic processing in a grid scalable environment, can be a valid help in the development of a new kind of Web 2.0 and new media applications to satisfy semantics processing and on-demand needs. Among end-user grid programming tools [12], we can cite, the Java Commodity Grid Toolkit (Java CoG Kit) that was created to assist in the development of applications using the well-known Globus Toolkit [11]; it is based on a workflow programming thanks to a XML language and the Karajan workflow engine. The GAUGE is another tool developed to work with Globus [8]. It generates full application code and allows users to focus on higher level abstraction avoiding low-level details.

1.1 Visual Processing for Media

Visual programming for media on Grid has to be able to formalize and represent concurrence of activities and the logic of services with an end-user oriented solution to simplify the development of complex applications. The visual tools have to be designed to help grid users to develop the application processes hiding the complexity and the technologies used (coding, access to databases, communications, coding format, parallel allocation, etc.). This kind of solutions are not only useful for Web 2.0/3.0 applications but also for many other massive applications.

In order to solve the above mentioned problems related to the visual programming tools for media on grid processing; a solution has been defined and validated on the AXMEDIS grid and model, starting from the AXMEDIS framework code of AXCP grid. The AXCP grid allows the formalization of processes for cross media content processing, semantic processing, content production, packaging, protection and distribution and much more [9], [10]. The work reported in this paper is related to the experience performed in defining a visual language for the formalization of visual media processing for Grid environments. The created visual model and tool can be adopted in other Grids as well.

2. AXCP Grid framework overview

The AXCP grid comprised of several Executors allocated to several processors for executing content processes is managed by a Scheduler. AXCP processes are called Rules, and are formalized in an extended JavaScript [9], [10]. The processes can be directly written in JS and/or the JS can be used to put in execution other processes. The Scheduler performs the rule firing, discovering grid Executors and managing possible problems/errors. The Scheduler (see Figure 1) may receive commands (to invoke sporadic or periodic rules with some parameters) and provide reporting information (e.g., notifications, exceptions, logs, etc…). The Scheduler exposes a Web Service which can be called by multiple applications such as web applications, WfMS, tools, and even other grid Rules on nodes of the AXCP.

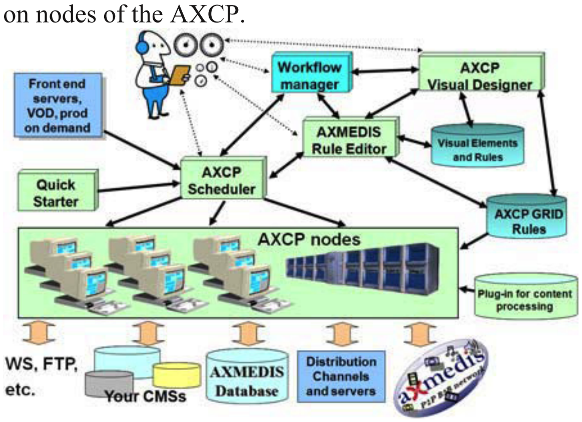


Figure 1 -- AXCP Architecture ©

The Executors receive the Rule to be executed from the Scheduler, and perform the initialization and the launch of the Rule on the processor. During the run, the Executor could send notifications, errors and output messages. Furthermore, the Executor could invoke the execution of other Rules sending a specific request to the Scheduler. This solution gains the advantages of a unified solution and allows enhancing the capabilities and the scalability of the AXMEDIS Content Processing [9], [10]. The AXCP processing tools are supported by a Plugin technology. The AXCP Rule language features allow to perform activities of ingestion, query and retrieval, storage, adaptation, extraction and processing descriptors, transcoding, synchronisation, fingerprint estimation, watermarking, indexing, summarization, metadata manipulation and mapping via XSLT, packaging, protection and licensing in MPEG-21 and OMA, publication and distribution via traditional channels and P2P.

AXCP Rules can be programmed by using the so called Rule Editor which is an IDE (Integrated Development Environment). The Rule editor is too technical to be used by non programmers, such as those that have to cope with the definition of content processing flow and activities in the content production factories.

3. The design of the Grid Visual Designer

Before starting with the development of the visual language for the AXCP, we performed a detailed analysis of all the AXCP rules developed and collected in the last three years by several users of AXCP (http://www.axmedis.org). From the analysis, it has been discovered that Rules collected and analyzed (about 280) were for the:

A. 75% single rules with a linear structure, presenting a sequence of activities to be performed. For each of them, when one of the activity fails the whole rule execution has to fail. To this category belong rules for automated content production on demand, licensing, content publication and/or repurposing, etc. These rules may have or may not have to report a result to the calling process which requested the execution of the Rule.
B. 9% rules activated by other Rules on the Grid in asynchronous manner. Their mother rule does not need to wait for the result to continue its running. These rules, even if they are activated by another Rule, are structurally realized as rules of type A. Since they start asynchronously and do not keep blocked the main rule as well.
C. 16% rules activating/invoking other processing Rules by creating synchronous/asynchronous derived Rules waiting/or-not for their completion to continue their execution.

Besides, we have discovered that almost all rules present JS segments of functional blocks working on single or on lists of content elements performing specific activities. This analysis allowed us to identify a possible semantics for a visual programming language based on composition of processing segment/blocks. Thus the Visual Program defined allows to compose:

- single elements of the process (called JSBlock) to create composed Rules allocated on the same processor node (covering rule of type A and B)
- branching activities (collection of RuleBlocks) which are allocated and executed on the Grid infrastructure according to their dependency by the scheduler. The Rules capable to activate other Rules cover the specific semantic of rules of type C, identified in the analysis.

The composition of these two models plus the implementation of a set of ready to use functional blocks (JSBlock or RuleBlocks) allowed to cover the issues mentioned in section 1.1 regarding: hierarchical structure, internal visual programming of single process flow on the single executing node, complex and branched flows composed by several different processes allocated on different grid nodes, error code reporting,, visual processing of media.

3.1 Modeling JSBlocks, Single Elements

According to the above reported analysis, a collection of visual blocks organized into a common repository and divided into categories has been created such as: Querying, ingesting, Posting, Metadata processing, Taking decision engine, Adapting/transcoding, Packaging, Licensing, AXMEDIS Object manipulation, Utility. In our visual programming model a generic block can be a segment of a JavaScript rule (a JSBlock) or a full RuleBlock (which in turn is created as a set of JSBlock or directly coded in JavaScript). JSBlocks can be composed by connecting inputs and outputs, according to their types; where each data value is an array that may contain a single element or a list of referred content or metadata elements. A JSBlock is characterized by a name (type name and instance) and a set of in/out parameters. A parameter can be marked as: (i) IN when it is consumed into the Rule, (ii) OUT when it is consumed into the rule and can be used to pass back a result to the next processing segments, that is IN/OUT, (iii) SETUP when it is a reserved INput to set up block specific behaviour. This parameter type is used to force different operating conditions in the Block. For example, to pass the ID of the database to be used, the temporary directory, etc.

Semantically speaking, a JSBlock is traduced into a JS Procedure specifying an elementary processing activity. Parameters are typed (String, Real, Integer and Boolean) and arrays of data have to be modelled as a string containing a list of formatted items by using specific separators. The JSBlock are connected one another according to their signature and arguments. The entry point function itself can invoke different functions defined in the same JavaScript (AXCP rule body). Skilled users may create their own Java Script JSBlocks augmenting their library.

3.2 Single Rule Visual Programming

JSBlocks can be combined to define the steps of a process flow corresponding to a grid processing rule, a RuleBlock (see Figure 2). The execution of JSBlocks is a sequential process flow according to the Boolean result returned by the previous one. Therefore, the process can take only one direction and end in one and only one of the leaves. The visual editor displays a green arrow for true and a red one for false. In this paper, we use a dashed arrow for false. The JSBlocks can be selected from the library of JSBlock including the functions listed in section 3.1. A single JSBlock can be quite complex; for example, it could activate other RuleBlocks which are processes on the grid, thus creating recursive or iterative patterns.

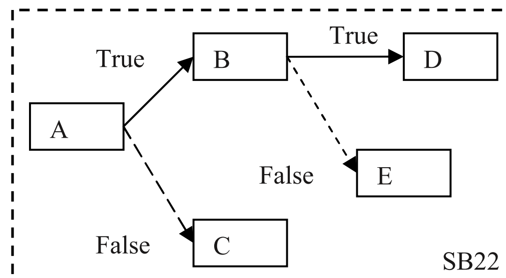


Figure 2 – A sequence of JSBlocks creating RuleBlock SB22

The visual programmer creates the specification with drag and drop approach, connecting blocks and imposing, through dialog boxes, the in/out parameters of a JSBlock (either by editing constant values or linking them with parameters of other blocks. In particular). JSBlocks composition is based on the forward and backward parameter propagation with the aim of creating a RuleBlock which is a rule to be allocated on a single node on the grid. The propagation allows linking the input parameters of a JSBlock with the IN/OUT parameters of its parents. In reference to the figure 2, input parameters of JSBlock D and E could be linked to IN/OUT parameters of B and A JSBlocks, whereas the JSBlock C sees only those of A. Backward propagation allows the definition of the IN/OUT parameters of the created RuleBlock by marking the input parameters of a JSBlock as global IN or OUT of the container RuleBlock.

Semantically speaking, the code generator starts by parsing the JSBlock sequence to produce a single RuleBlock by assembling all the JSBlocks in a single

RuleBlock, including the JS code and the maps of parameters among JSBlocks. Finally the IN and OUT parameter definitions are created and assigned to the signature of the RuleBlock implementing the calls chain. Finally, the resulting RuleBlock, JS Rule, is activated on the grid according to its parameters.

At a first glance, the visual programming semantics of RuleBlocks **seems** to have relevant limitations, but it is not fully true, as put in evidence in the following. In fact, it should be considered that on the basis of the above described model the: (i) **iterations** are internally managed into the single JSBlocks. (ii) **decisions** can be taken into the single JSBlock. A JSBlock can be regarded as a visual implementation of a selection and/or of a sequence of actions. On the other hand, the above semantics does not address the modeling of multiple branches into the graphs, and thus the management of multiple rules/processes.

3.3 Managing RuleBlocks Rules on grid

The visual programming model proposed has a specific modality to specify branched activations of rules on the grid, by delegating to the grid scheduler the effective allocation of processes on nodes. To this purpose, a different visual model/semantics, with respect to previous one, has been defined. It allows the construction of branched and distributed rules. In this case, see Figure 3, the visual graph is a tree represents a set of processes and their activation relationships, as depicted in Figure 3.

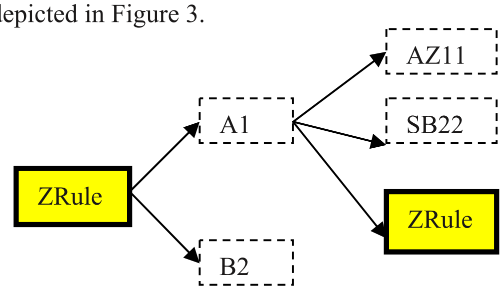


Figure 3 – ZRule ManRuleBlock defined to manage Rules on grid.

A Managing RuleBlock (ManRuleBlock) ZRule is created, for example, to activate the execution of RuleBlocks A1 and B2 in a sequential way (or parallel) on the grid and to return on the process ZRule synchronously (or asynchronously) their return parameters. Even in this case, the single RuleBlock can be selected from a library or can be created by using

- visual model of section 3.2, see figure 2;
- AXCP Editor for JavaScript as single JS Rule to be used as a block in the library of rules;
- Another ManRuleBlock, defined by using Visual programming model depicted in Figure 3 (for example, one of the child of ZRule is an instance of the ZRule, invoked by A1 with some parameters.

Regarding IN/OUT parameter management and editing, a ManRuleBlock follows the same semantics of the RuleBlock thanks to the forward and backward parameter propagation. Please note that the definition of ManRuleBlock can be recursive as depicted in Figure 3, in which ZRule is calling via A1 another instance of ZRule.

Semantically speaking, the code generator produces the a JS Rule implementing the ManRuleBlock (e.g, ZRule in example of Figure 3) for managing the activation of other RuleBlocks according to the graph and always respecting the assignment of parameters of the RuleBlocks. Please note that RuleBlocks are activated by using a web service calls of the Scheduler. The code generator produces the a Rule called ZRule which is the invoker and also the manager for IN/OUT parameters, waiting for the answers/results of the called RuleBlocks to pass them to the others according to the flow.

4. Example of visual programming grid

In this section, an example of visual programming is reported. It implements an audio recognition process based on fingerprint, to recognize audio track on the basis of audio fingerprint, for examples when the audio are uploaded on a portal to filter out User Generated Content, UGC, infringing copyrights.

The first step has been the identification of basic procedures involved in audio file searching, fingerprint extraction, database insertion and searching. The JSBlocks have been used to compose and generate RuleBlocks. Combining them, by linking parameters, the following RuleBlocks were built:

A) FingerprintExtraction (in::folderIn, in::fileExt, in::folderOut). It uses getFileList, extractFingerprint and alert Jsblocks (see Figure 4). The *fileList* input parameter of "fingerprint_extraction" procedure is associated with the *filePathList* out parameter of "fileList". The folderIn, fileExt (getFileList) and folderOut (extractFingerprint) are back-propagated as input parameter for the "FingerprintExtraction" RuleBlock.

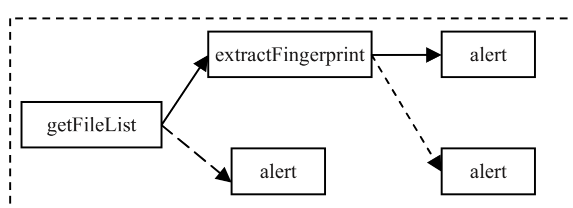


Figure 4 – RuleBlock FingerprintExtraction

B) InsertIntoDB (in::folderIn, in::fileExt)

This RuleBlock uses getFileList, insertFingerprint, alert Jsblocks (see Figure 5). The *fileList* input parameter of "insertFingerprint" procedure is associated with the *filePathList* out parameter of "getFileList".

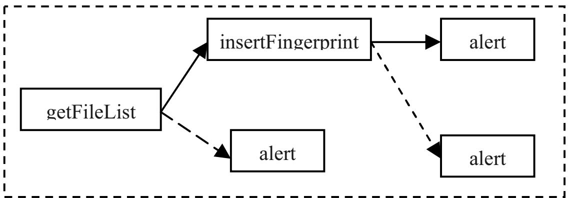


Figure 5 – RuleBlock InsertIntoDB

The *folderIn* and *fileExt* (getFileList) are back-propagated as input parameter for the "InsertIntoDB" RuleBlock.

C) SearchIntoDB (in::folderIn, in::fileExt, in::resultFilePath, in::dbID).

It uses getFileList, searchFingerprint and alert Jsblocks (see Figure 6). The *fileList* input parameter of "searchFingerprint" procedure is associated with the *filePathList* out parameter of "getFileList". The *folderIn*, *fileExt* (getFileList), *resultFilePath* and *dbID* (searchFingerprint) are back-propagated as input parameters for the "SearchIntoDB" RuleBlock.

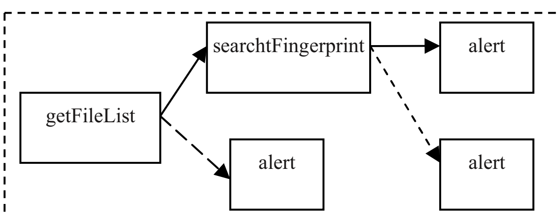


Figure 6 – RuleBlock InsertIntoDB

By using the ManBlockRule Visual Programming RuleBlocks were used to build the more complex RuleBlocks.

AddNewFingerprint is put in execution every time new audio files are added in the repository, then it extracts fingerprints, and inserts them into database.

This new RuleBlock uses the following:

- FingerprintExtraction(in::folderIn, in::fileExt, in::;folderOut)
- InsertIntoDB(in::folderIn, in::fileExt)

The *folderIn*, *fileExt* and *folderOut* (FingerprintExtraction) are back-propagated as input parameters. The *folderIn* input of "InsertIntoDB" rule and the *folderOut* input of "FingerprintExtraction" are both associated with the *folderOut* parameters of the new RuleBlock. The *fileExt* of "FingerprintExtraction" is associated with the *fileExt* of new RuleBlock in order to define the wildcard for audio files to get all the files stored in the *folderIn*.

SearchFingerprintInDatabaset RuleBlock runs when a unknown audio file has to be identified by searching its fingerprint inside the database.

This new RuleBlock uses (see Figure 7):

- FingerprintExtraction(in::folderIn, in::fileExt, in::;folderOut)
- SearchIntoDB(in::folderIn, in::fileExt, in::resultFilePath, in::dbID)
-

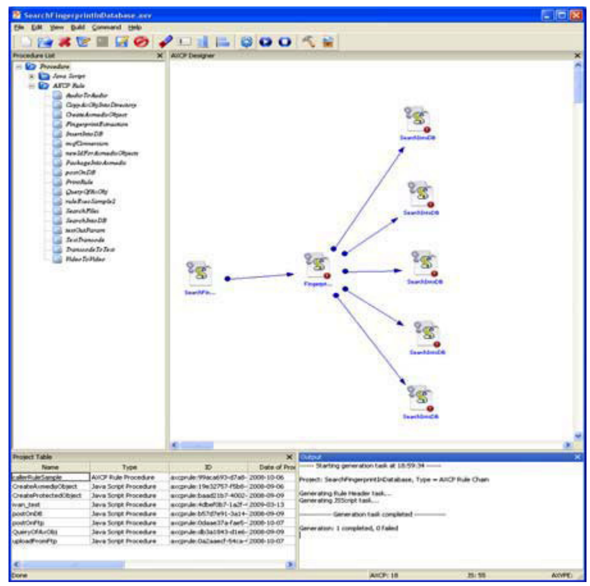


Figure 7 – ManRuleBlock SearchFingerprintInDatabase

The presented Rule Blocks were used inside the AXMEDIS GRID environment to populate a fingerprint database starting from a large collection of audio tracks. The fingerprint extraction algorithm works on mp3 and wave audio formats normalizing the audio features (sample rate, number of channels and bit per sample) when necessary and generates a fingerprint by using the Spectral Flatness descriptor.

The production of the experiments have been quite fast and simple for the visual programmer. Several other ManRuleBlocks have been defined to replicated via visual programming real grid rules provided by partners.

The visual language and semantic model proposed allowed to cope with almost all grid patterns identified, and ranging from sequential to parallel execution, asynchronous and synchronous invocations, recursive and iterative, etc. The approach resulted quite effective and usable. It has been strongly appreciated since the users can customize the single JSBlock, and may create their own Blocks according to the specific application domain in which they have to work. Additional work for the modeling is needed to manage versioning of the visual elements and to allow the semantic search of the Blocks into the database of reusable blocks. Presently, the search of the most suitable blocks is supported by a table that represents all the main features crossing the input and outputs parameters with the main data item processed by the single block. We have noticed that the reuse of blocks is mainly performed at level of JSBlocks. ManRuleBlocks and RuleBlocks are quite frequently versioned, adding more parameters and pushing them to become more general to be reused in several occasions, this can be a problem since the same Block can be used by several

5. Conclusions

In this paper, a visual programming model for content processing grid has been proposed. It has been designed for general grid processing and implemented for validation on AXCP grid open solution. The featured have been identified by analysing a large set of real processing grid rules. The derived model satisfied the 97% of them. On the other hand, the code generator allows to access the code adjusting the uncovered missing 3% of rules. A remodelled rules have been tested against the optimized manually create rules. In most cases the rule visually produced present lower performance of the original one. A further work is in process to add some more constructs that may enables the visual programmer to manage with more simplicity the errors recovering in the ManRuleBlock semantics and in defining rules that are activated by multiple firing conditions. Presently these issues are managed at the level of code with specific JSBlocks while in some cases they are constructs that should be visible at the higher level of grid programming. The full documentation can be recovered on the AXMEDIS portal http://www.axmedis.org.

Acknowledgments

The authors would like to thank all the AXMEDIS partners for their contributions. Most of the work reported has not been performed after the AXMEDIS project completion and offered to the framework that is still growing.

References

[1] Zhiwei Xu, Chengchun Shu, Haiyan Yu, Haozhi Liu, “An Agile Programming Model for Grid End Users” Proceedings of the Sixth Int. Conf. on Parallel and Distributed Computing, App. and Tech. (PDCAT’05)

[2] Special Issue: Grid Computing Environments. *Concurrency and Computation: Practice and Experience*, 14:1035-1593,2002.

[3] J. Novotny. The grid portal development kit. *Concurrency and Computation: Practice and Experience*, 14:1129-1144, 2002.

[4] E. Akarsu, F. Fox, W. Furmanski, and T. Haupt. WebFlow – high level programming environment and visual authoring toolkit for high performance distributed computing. In *Proceedings of the 1998 ACM/IEEE Conference on Supercomputing*, pages 1-7, 1998.

[5] R. Armstrong, D. Gannon, A. Geist, K. Keahey, S. Kohn, L. McInnes, S. Parker, and B. Smolinski. Toward a common component architecture for high performance scientific computing. In *Proc. Of 8th. IEEE Int. Symp. on High Performance Distributed Computing,* 1999.

[6] E. Deelman, J. Blythe, Y. Gil, C. Kesselman, G. Mehta, K. Vahi, K. Blackburn, A. Lazzarini, A. Arbree, R. Cavanaugh, and S. Koranda. Mapping abstract complex workflows onto grid environments. *Journal of Grid Computing*, 1:25-39,2003.

[7] M. Lorch and D. Kafura. Symphony – A java-based composition and manipulation framework for computational grids. In *Proc. of the 2nd IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGrid2002)*, pages 136-143, 2002.

[8] F. Hernández, P. Bangalore, J. Gray, Z. Guan, and K. Reilly. GAUGE: Grid automation and generative environment. *Concurrency and Computation: Practice and Experience*, to appear. 2005.

[9] P. Bellini, I. Bruno and P. Nesi, “A Distributed Environment for Automatic Multimedia Content Production based on GRID”, in Proc. AXMEDIS 2005, Florence, Italy, 30/11-2/12, pp134-142, IEEE Press.

[10] P. Bellini, I. Bruno, P. Nesi, ``A language and architecture for automating multimedia content production on grid", Proc. of the IEEE International Conference on Multimedia & Expo (ICME 2006), IEEE Press, Toronto, Canada, 9-12 July, 2006.

[11] F. Hernández, P. Bangalore, J. Gray, and K. Reilly. A graphical modeling environment for the generation of workflows for the globus toolkit. In V. Getov and T. Kielman, editors, *Component Models and Systems for Grid Applications. Proceedings of the Workshop on Component Models and Systems for Grid Applications held June 26, 2004 in Saint Malo, France*, pages 79-96. Springer, 2005.

[12] Minglu Li, Jun Ni, Qianni Deng, Xian-He Sun, “Grid and cooperative computing”: second international workshop, GCC 2003, Shanhai [sic], China, December 7-10 2003 : revised papers, Springer, 2004, ISBN 3540219889, 9783540219880

Interactive Multimedia Systems for Technology-Enhanced Learning and Preservation

Kia Ng,[1] Eleni Mikroyannidi,[1] Bee Ong,[1] Nicolas Esposito[2] and David Giaretta[3]
[1]ICSRiM - University of Leeds, School of Computing & School of Music, Leeds LS2 9JT, UK
[2]Costech and Heudiasyc Labs, University of Technology of Compiègne and CNRS, Centre Pierre Guillaumat, 60200 Compiègne, France
[3]STFC, Rutherford Appleton Laboratory, Oxfordshire OX11 0QX, UK
info@icsrim.org.uk www.icsrim.org.uk

Abstract

Interactive multimedia and human-computer interaction technologies are effecting and contributing towards a wide range of developments in all subject areas including contemporary performing arts. These include augmented instruments for interactive music performance, installation arts and technology-enhanced learning. Consequently, the preservation of interactive multimedia systems and performances is becoming important to ensure future re-performances as well as preserving the artistic style and heritage of the art form. This paper presents two interactive multimedia projects for technology-enhanced learning, and discusses their preservation issues with an approach that is currently being developed by the CASPAR EC IST project.

Keywords: Interactive Multimedia Performance, Technology-enhanced learning, Motion capture, sensor, multimodal, Digital Preservation, Ontologies.

1. Introduction

Interactive multimedia technologies and all forms of digital media are popularly used in contemporary performing arts, including musical compositions, installation arts, dance, etc. Typically, an Interactive Multimedia Performance (IMP) involves one or more performers who interact with a computer based multimedia system making use of multimedia content. This content may be prepared and generated in real-time and may include music, manipulated sound, animation, video, graphics, etc. The interactions between the performer(s) and the multimedia system can be done in a wide range of different approaches, such as body motions [1, 2], movements of traditional musical instruments, sounds generated by these instruments [3, 4] , tension of body muscle using bio-feedback [5], heart beats, sensors systems, and many others. These "signals" from performers are captured and processed by multimedia systems. Depending on specific performances, the "signals" will be mapped to multimedia content for generation using a mapping strategy (see Figure 1).

An example of an IMP process is the one adopted in the MvM (Music via Motion) interactive performance system, which produces music by capturing user motions [1, 6].

Interactive multimedia systems have been applied in a wide range of applications in this context. This paper presents two interactive multimedia systems that are designed for technology-enhanced learning for music performance (one for string instruments playing and one for conducting) and interactive multimedia performance, and consider their preservation issues and complexity.

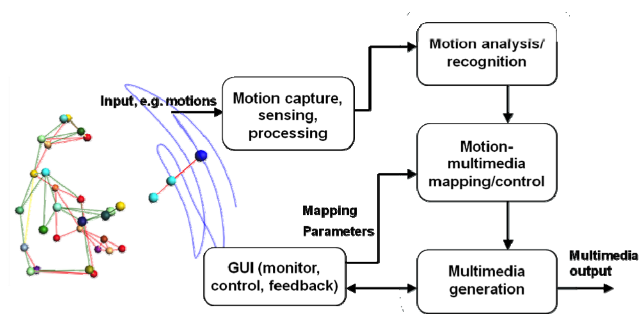


Figure 1: Interactive Multimedia Performance process

Generally, manipulating/recording multimedia content using computers is an essential part of a live interactive performance. Using simply performance outputs recorded in the form of audio and video media will not be sufficient for a proper analysis (e.g. for studying the effect of a particular performing gesture on the overall quality of the performance) or reconstruction of a performance at a later time. In this context, traditional music notation as an abstract representation of a performance is also not sufficient to store all the information and data required to reconstruct the performance. Therefore, in order to keep a performance alive through time, not only its output, but also the whole production process to create the output needs to be preserved.

The remaining paper is organized as follows. Section 2 presents two Interactive Multimedia Performance Systems that need to be preserved. Section 3 introduces the conceptual model of the CASPAR project and the tools that are used for the preservation of the IMP systems. Finally the paper is concluded in section 4 and the next steps of future work are outlined.

2. Interactive Multimedia Performance Systems (IMP)

2.1. 3D Augmented Mirror (AMIR)

The 3D Augmented Mirror (AMIR) [7, 8, 9] is an IMP system being developed in the context of the i-Maestro (www.i-maestro.org) project, for the analysis of gesture and posture in string practice training. String players often use mirrors to observe themselves practicing. More recently, video has also been used. However, this is generally not effective due to the inherent limitations of 2D perspective views of the media.

Playing an instrument is physical and requires careful coaching and training on the way a player positions himself/herself with the aim to provide the best/effective output with economical input, i.e. least physical effort. In many ways, this can be studied with respect to sport sciences to enhance performance and to reduce self inflicted injuries.

With the use of 3D Motion Capture technology, it is possible to enhance this practice by online and offline visualising of the instrument and the performer in a 3D environment together with precise and accurate motion analysis to offer a more informed environment to the user for further self-awareness, and computer assisted monitoring and analysis.

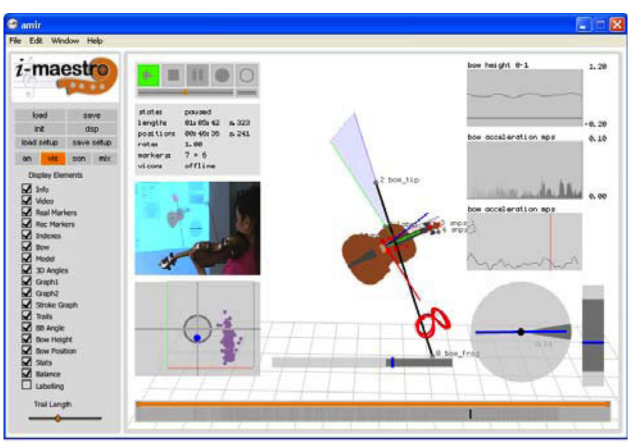


Figure 2: Graphical Interface of the 3D Augmented Mirror System

The 3D Augmented Mirror is designed to support the teaching and learning of bowing technique, by providing multimodal feedback based on real-time analysis of 3D motion capture data. Figure 2 shows a screenshot of the 3D Augmented Mirror interface, including synchronized video and motion capture data with 3D bowing trajectories.

When practicing using AMIR, a student can view the posture and gesture sequences (3D rendering of the recorded motion data) as prepared by the teacher, selecting viewpoints and studying the recording without the limitations of a normal 2D video. A student can also make use of the system to capture and study their own posture and gesture, or to compare them with some selected models.

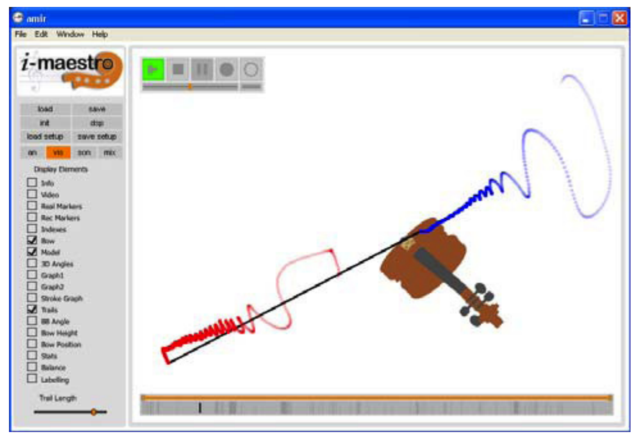


Figure 3: Gesture *signature* - tracing gesture for the analysis of composition.

It has been found that the AMIR multimodal recording which includes 3D motion data, audio, video and other optional sensor data (e.g. balance, etc) can be very useful to provide in-depth information beyond the classical audio visual recording for musicological analysis (see Figure 3). Preservation of the IMP system is of great importance in order to allow future re-performance. The multimodal recoding offers an additional level of detail for the preservation of musical gesture and performance that can be vital for the musicologist of the future. These contributions have resulted in our motivation for the preservation of the AMIR multimodal recordings.

2.2. ICSRiM Conducting Interface

The ICSRiM Conducting System is another IMP system developed for the tracking and analysis of a conductor's hand movements [10, 11]. Its aim is to help students learning and practicing conducting.

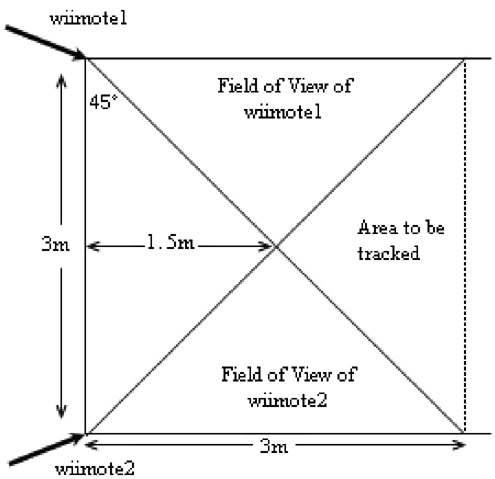


Figure 4: Wii-based 3D capture setup.

A portable motion capture system composed by multiple Nintendo Wiimotes is used to capture the conductor's gesture. The Nintendo Wiimote has several advantages as it combines both optical and sensor based motion tracking capabilities, it is portable, affordable and easily attainable. The captured data are analyzed and presented to the user in an entertaining as well as pedagogically informed manner highlighting important factors and offer helpful and informative monitoring for raising self awareness that can be used during a lesson or for self-practice. Figure 5 shows a screenshot of the Conducting System Interface with one of the four main visualization mode.

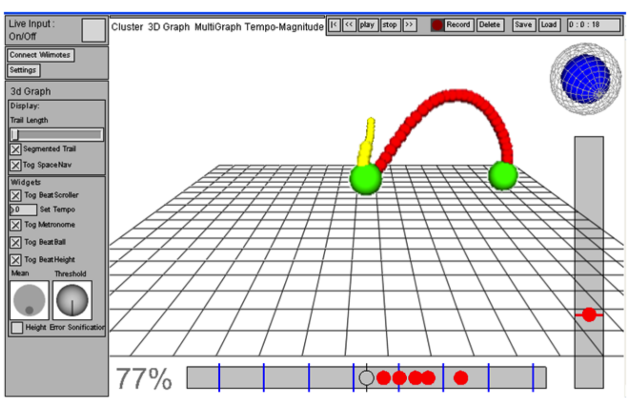


Figure 5: Graphical Interface of the ICSRiM Conducting System.

3. Preservation

Preserving the whole production process of an IMP is a challenging issue. In addition to the output multimedia contents, related digital contents such as mapping strategies, processing software and intermediate data created during the production process (e.g. data translated from "signals" captured) have to be preserved, together with all the configuration, setting of the software, changes (and time), etc. Both Multimedia Systems presented on section 2, generate similar type of datasets. The dataset usually consists of the captured 3D motion data, video and audio files, MAX/MSP patches and additional configuration files. The reproduction of the IMP can be achieved through the correct connection of these components. Therefore, the most challenging problem is to preserve the knowledge about the logical and temporal relationships among these individual components so that they can be properly assembled into a performance during the reconstruction process.

Another important aspect that needs to be preserved is also the comments and feedbacks that are generated from the users or performer during the production of an IMP and regard the quality of the performance and the used techniques. In the context of the CASPAR project, we have adopted an ontology-driven approach [13-15] that reuses and extends existing standards, such as the CIDOC Conceptual Reference Model (CIDOC-CRM) [16, 17] for the efficient preservation of an IMP.

3.1. Conceptual Model of CASPAR Preservation

The CASPAR framework is based on the full use of the OAIS (Open Archival Information System) Reference Model [18], which is an ISO standard. The OAIS conceptual model is shown in Figure 6. The Conceptual Model aims to provide an overall view of the way in which the project sees preservation working. Also the conceptual model helps to highlight the areas which can help to the formation of an interoperable and applicable structure that can support effectively the digital preservation across the different CASPAR domains.

The very basic concept defined in the OAIS Reference Model is the Information Object. As illustrated in the UML diagram of Figure 6, an Information Object is composed of a Data Object and one or more layers of Representation Information. A Data Object can be a Physical Object (e.g. a painting) or a Digital Object (e.g. a JPEG image). Representation Information provides the necessary details for the interpretation of the bits contained within the digital object into meaningful information. For digital objects, representation information can be documentation about data formats and structures, the relationships amongst different data components. Representation information can also be software applications that are used to render or read the digital objects.

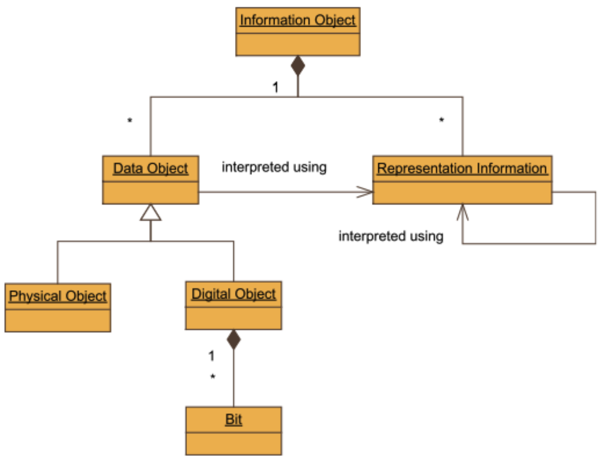


Figure 6: Basic concepts of OAIS Reference Model - Information Object [18]

In addition, the Representation needs to be connected with the Knowledge base of the designated community. Ontology models offer the means for organizing and representing the semantics of this knowledge base.

3.2 The ICSRiM Archival System

The Archival System has been developed by the University of Leeds and it is used for the access, retrieval and preservation of different IMPs. The architecture of the Archival system is based on the OAIS conceptual model and on the CASPAR Framework. In addition, the Archival system integrates the appropriate CASPAR components (http://www.casparpreserves.eu/publications/software-releases/) as web services for the efficient preservation of the IMP.

The architecture of the Archival system is shown in Figure 7. It has been designed in order to support the preservation of different types of IMPs. Thus, it can be used for both the 3D Augmented Mirror and the Conducting System.

The archival system provides a web interface and its backend communicates with a Repository containing the IMPs and the necessary metadata for preserving the IMPs. Before the ingestion of an IMP, it is necessary to create its description based on the CIDOC-CRM and FRBRoo ontologies. This information is generated in RDF/XML format with the use of the CASPAR Cyclops tool (http://www.utc.fr/caspar/wiki/pmwiki.php?n=Main.Proto). The Cyclops tool is used to capture appropriate Representation Information from a high level in order to enhance virtualization and future re-use of the IMP. It also offers the ability of adding comments and annotations concerning any concept of the IMP. Figure 8 shows the Graphical interface of the Cyclops tool and how it is used to create an IMP description. The tool provides a palette for creating the description of an IMP as a graph in the drawing area.

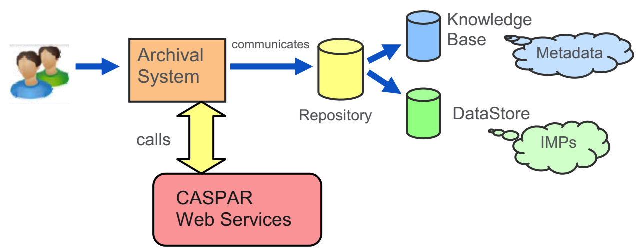


Figure 7: The Architecture of the ICSRiM Archival System.

Figure 9 shows in detail the graphical instantiation of an IMP that was created with the use of the Cyclops tool. The graph can capture information about the software and hardware that was used as well as the components that were produced (e.g. 3D motion data) and how they are linked these components for the reproduction of an IMP. The concepts of the diagram shown in Figure 9 can be mapped to the concepts of the CIDOC-CRM and FRBR ontologies**Error! Reference source not found.**. However, the usable interface of the tool hides the complexity of the system from the user. It uses a simple high level language (concepts, relations, and types) which is based on the terminology of the domain and does not require any ontology expertise to create the instantiation. The Cyclops canvas offers a graphical representation of the life cycle to make its understanding easier. Furthermore, Cyclops is a Web application, facilitating the portability. It is open source and it uses the following technologies: XUL, JavaScript, SVG, HTML, CSS, XML, PHP, MySQL. Cyclops can be used as an integrated component of the ICSRiM Archival System as well as a standalone application.

The retrieval of an IMP is based on queries that are applied on the Knowledge Base. In particular, the Web Archival calls the FindingAids services, which task is to perform RQL queries on the Representation Information Objects and return the results to the user. Every Representation Information object is linked to a corresponding dataset of an IMP stored in the Repository.

Therefore, the user will be able to retrieve the IMP files s/he is interested in and their description.

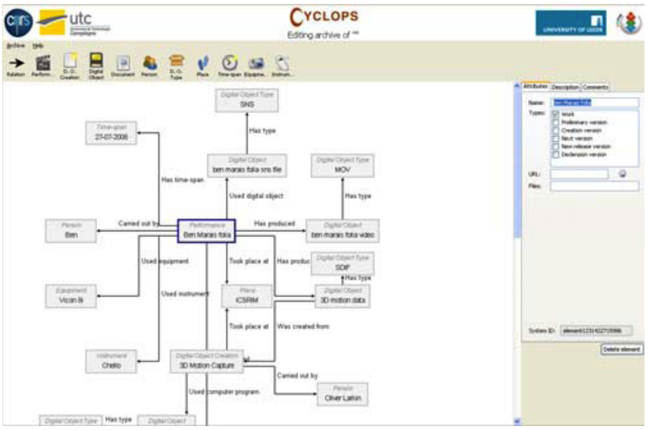


Figure 8: The graphical interface of the Cyclops tool.

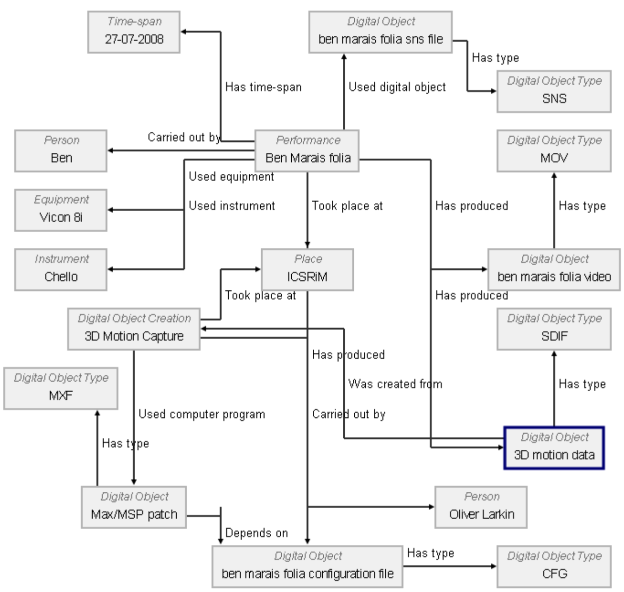


Figure 9: An IMP instantiation created with the Cyclops tool.

4. Conclusions and Future Work

The paper presented the CASPAR Conceptual model and the tools that are used for the preservation of interactive multimedia performances. The approach of the project considers ontologies as a semantic knowledge base containing the necessary metadata for the preservation of IMPs.

The design of the system offers flexibility in preserving multiple IMP systems. In addition, the preservation of the IMP Systems could enhance the learning procedure as it provides ways of capturing feedbacks and comments on the quality of the IMP. It also helps to preserve the intangible heritage that an IMP reflects.

We are currently working on the deployment of the CASPAR components within the Archival System. In particular, we are integrating software tools such as the Semantic Web Knowledge Middleware [19], for performing Information Retrieval tasks that will facilitate the exploitation of our knowledge base.

5. Acknowledgements

Work partially supported by European Community under the Information Society Technologies (IST) programme of the 6th FP for RTD - project CASPAR. The authors are solely responsible for the content of this paper. It does not represent the opinion of the European Community, and the European Community is not responsible for any use that might be made of data appearing therein.

The research is supported in part by the European Commission under Contract IST-026883 I-MAESTRO. The authors would like to acknowledge the EC IST FP6 for the partial funding of the I-MAESTRO project (www.i-maestro.org), and to express gratitude to all I-MAESTRO project partners and participants, for their interests, contributions and collaborations.

The authors would also like to acknowledge David Bradshaw for his work on the development of the ICSRiM Conducting Interface.

6. References

[1] K. C. Ng, "Music via Motion: Transdomain Mapping of Motion and Sound for Interactive Performances," *Proceedings of the IEEE,* vol. 92, 2004.

[2] R. Morales-Manzanares, E. F. Morales, R. B. Dannenberg, and J. Berger, "SICIB: An Interactive Music Composition System Using Body Movements," *Computer Music Journal,* vol. 25, pp. 25-36, 2001.

[3] D. Young, P. Nunn, and A. Vassiliev, "Composing for Hyperbow: A Collaboration between MIT and the Royal Academy of Music," in *International Conference on New Interfaces for Musical Expression*, Paris, France, 2006.

[4] D. Overholt, "The Overtone Violin," in *International Conference on New Interfaces for Musical Expression*, Vancouver, BC, Canada, 2005.

[5] Y. Nagashima, "Bio-Sensing Systems and Bio-Feedback Systems for Interactive Media Arts," in

2003 Conference on New Interfaces for Musical Expression (NIME-03), Montreal, Canada, 2003.
[6] MvM, "Music via Motion.", http://www.icsrim.org.uk/mvm
[7] K. Ng, Technology-Enhanced Learning for Music with i-Maestro Framework and Tools, in Proceedings of EVA London 2008: the International Conference of Electronic Visualisation and the Arts, British Computer Society, 5 Southampton Street, London WC2E 7HA, UK, ISBN: 978-1-906124-07-6, 22-24 July 2008.
[8] K. Ng, Interactive Feedbacks with Visualisation and Sonification for Technology-Enhanced Learning for Music Performance, in Proceedings of the 26th ACM International Conference on Design of Communication, SIGDOC 2008, Lisboa, Portugal, 22-24 September 2008.
[9] K. Ng, O. Larkin, T. Koerselman, B. Ong, D. Schwarz, and F. Bevilaqua., "The 3D Augmented Mirror: Motion Analysis for String Practice Training," in Proceedings of the *International Computer Music Conference (ICMC2007)*, Copenhagen, Denmark, 2007.
[10] D. Bradshaw, and K. Ng, Tracking Conductors Hand Movements using Multiple Wiimotes, in Proceedings of the International Conference on Automated Solutions for Cross Media Content and Multi-channel Distribution (AXMEDIS 2008), 17-19 Nov. 2008, Florence, Italy, pp. 93-99, Digital Object Identifier 10.1109/AXMEDIS.2008.40, IEEE Computer Society Press, ISBN: 978-0-7695-3406-0. 4.
[11] D. Bradshaw and K. Ng, Analyzing a Conductor's Gestures with the Wiimote, in Proceedings of EVA London 2008: the International Conference of Electronic Visualisation and the Arts, British Computer Society, 5 Southampton Street, London WC2E 7HA, UK, 22-24 July 2008.
[12] K. Ng, T.V. Pham, B. Ong, A. Mikroyannidis, D. Giaretta, Preservation of interactive multimedia performances, International Journal of Metadata, Semantics and Ontologies 2008 - Vol. 3, No.3 pp. 183 – 196, 10.1504/IJMSO.2008.023567
[13] K. Ng, A. Mikroyannidis, B. Ong, D. Giaretta, Practicing Ontology Modelling for Preservation of Interactive Multimedia Performances, in Proceedings of the International Conference on Automated Solutions for Cross Media Content and Multi-channel Distribution (AXMEDIS 2008), 17-19 Nov. 2008, Florence, Italy, pp. 276-281, Digital Object Identifier 10.1109/AXMEDIS.2008.43, IEEE Computer Society Press, ISBN: 978-0-7695-3406-0.
[14] K. Ng, T. V. Pham, B. Ong, A. Mikroyannidis, and D. Giaretta, "Ontology for Preservation of Interactive Multimedia Performances," in *2nd International Conference on Metadata and Semantics Research (MTSR 2007)*, Corfu, Greece, 2007.
[15] A. Mikroyannidis, B. Ong, K. Ng, and D. Giaretta, "Ontology-Based Temporal Modelling of Provenance Information," in *14th IEEE Mediterranean Electrotechnical Conference (MELECON'2008)*, Ajaccio, France, 2008.
[16] T. Gill, "Building semantic bridges between museums, libraries and archives: The CIDOC Conceptual Reference Model," *First Monday,* vol. 9, 2004.
[17] M. Doerr, "The CIDOC CRM - an Ontological Approach to Semantic Interoperability of Metadata," *AI Magazine,* vol. 24, 2003.
[18] Consultative Committee for Space Data Systems, "Reference Model for An Open Archival Information System," 2002.
[19] D. Zeginis, Y. Tzitzikas, and V. Christophides, "On the Foundations of Computing Deltas Between RDF Models," in *6th International Semantic Web Conference (ISWC-07)*, Busan, Korea, 2007.

LoCa – Towards a Context-aware Infrastructure for eHealth Applications*

Nadine Fröhlich[1] Andreas Meier[2] Thorsten Möller[1]
Marco Savini[2] Heiko Schuldt[1] Joël Vogt[2]

[1] Department of Computer Science, University of Basel, Switzerland
[2] Department of Informatics, University of Fribourg, Switzerland

E-mail: {nadine.froehlich, thorsten.moeller, heiko.schuldt}@unibas.ch
{andreas.meier, marco.savini, joel.vogt}@unifr.ch

Abstract

New sensor technologies, powerful mobile devices and wearable computers in conjunction with wireless communication standards have opened new possibilities in providing customized software solutions for medical professionals and patients. Today, medical professionals are usually equipped with much more powerful hardware and software than some years before. The same is true for patients which, by making use of smart sensors and mobile devices for gathering, processing and analyzing data, can live independently in their home environment while receiving the degree of monitoring they would get in stationary care. All these environments are highly dynamic, due to the inherent mobility of users. Therefore, it is of utmost importance to automatically adapt the underlying IT environment to the current needs of their users – which might change over time when user context evolves. In a digital home environment, this requires the automatic customization of user interfaces and the context-aware adaptation of monitoring workflows for mobile patients. This paper introduces the LoCa project which will provide a generic software infrastructure, able to dynamically adapt user interfaces and services-based distributed applications (workflows) to the actual context of a user (physician, caregiver, patient). In this paper, we focus on the application of LoCa to monitoring the health state of mobile patients in a digital home environment.

1. Introduction

Telemonitoring applications enable healthcare institutions to control therapies of patients in out-of-hospital settings. In particular, telemonitoring allows patients to live as independently as possible in their digital home environment. The goal is to support the individual disease management by patient monitoring which will result in less hospitalization and a higher quality of life. In the presence of an increasingly aging population and a growing number of people suffering from chronic ailments, this kind of applications already has a high relevance for the healthcare system and is expected to gain even more importance.

Monitoring includes the continuous gathering, processing and analysis of mainly physiological data coming from sensors which are either integrated into the patient's digital home or attached to the patient's body or clothes. Currently, these monitoring applications are rarely automated. Configuration of the sensor environment, the customization for a particular patient, and the actual data processing and analysis are mostly tedious manual tasks. In the LoCa project (A Location and Context-aware eHealth Infrastructure), we aim at providing a user-friendly and adaptable solution for the automated gathering and analysis of relevant data for monitoring patients. LoCa will be a general purpose system that can be applied both in digital home environments and in stationary care. A main feature in LoCa is the consideration of context as a first class citizen. This means that monitoring applications and processes as well as user interfaces will be dynamically adapted based on the user's context (e.g., location, activity, etc.). Context-aware adaptations will result in more customized monitoring solutions and thus better support for data analysis and emergency assistance (e.g., triggering of emergency services in case of severe health conditions). Dynamic adaptations will also allow to seamlessly apply best practices in health monitoring and patient control without explicit reconfigurations.

Consider, for instance, a sixty-five year old male patient with cardiac problems in convalescence. During his recovery at home, his physician would like to control his state of health and therefore needs to continuously receive data on his physiological condition. At the moment, the patient's ECG is measured periodically once a day or additionally, in

*The LoCa project is funded by the Hasler Foundation.

case the patient does not feel well. For this, a nurse is sent to the patient's home to record ECG data and other measurements. The physician only receives raw data and has to manually initiate all the steps needed for the interpretation of raw data in a particular order, including a comparison of the actual values with the patient's medical history, to determine the individual development of physiological data.

In order to improve this situation, the patient is given a smart shirt equipped with several sensors metering physiological parameters like ECG and blood glucose level. In addition, the patient receives a smart phone with GPS sensor and camera. From the point of view of the patient, this allows for almost unlimited mobility and does no longer require him to stay at home for the necessary measurements. From the physician's point of view, the smart shirt allows for the continuous gathering of vital parameters and thus for seamless monitoring in real time. As an important requirement for properly analyzing and interpreting metered data, the physician needs to know the exact context of the measurement (e.g., the patient's location and activity). Therefore, the shirt not only has to provide physiological data but also details on his activity (e.g., by means of acceleration sensors that can monitor the physical exercises he is doing). The patient's therapy includes a healthy diet, without alcohol and cigarettes, as well as physical exercises he is not used to. Thus, he writes an electronic diary, extended with photos of his meals, which finally helps in communicating diet information and stress factors to his physician. Annotations to this diary, provided by the physician, support the patient in understanding effects of his behavior for his therapy. Having access to raw sensor data does not yet allow the physician to properly analyze the patient's health state. The data still has to be cleaned, eventually coarsened, and analyzed in correlation with each other. For data analysis, the physician will follow a process consisting or dedicated processing steps in pre-defined order. To ease her work she will use the LoCa system to define these workflows in a user-friendly way, thereby determining rules for data interpretation. Finally, she is able to define proper thresholds, for instance for critical blood pressure values in stress situations.In case a threshold is exceeded, the physician will be visually advised on her screen or will receive an SMS. It is important to note that neither the analysis processes nor the corresponding user interfaces are static but need to be automatically adapted as soon as the context of the patient changes (e.g., when a different set of sensors is available), or in the course of the therapy when further parameters need to be taken into account.

The objective of the LoCa project is to address the challenges introduced above and provide reliable support for workflow-based eHealth applications. This includes telemonitoring in home care as well as applications in stationary care. In close collaboration with stakeholders from the healthcare domain, different use cases from both applications have already been defined. Finally, the LoCa system will be applied and evaluated in a stationary care and in a home care environment by the medical project partners. In this paper, we focus on telemonitoring applications in a digital home environment. From a functional perspective, the goal is to gather, process, analyze, and visualize physiological data and to store aggregated data in the electronic health record of a patient. In particular, the analysis and visualization will be dynamically tailored to the patient's context. This includes sophisticated failure handling which, by considering context at run-time, does not need to be prespecified in monitoring workflows. The system should finally be able to detect and anticipate potential cardiac irregularities or other health-related problems, based on criteria defined by the medical partners in the project. From a systems point of view, LoCa will make use and extend an existing platform for the reliable processing of data streams for health monitoring across fixed and mobile devices [5, 6].

In this paper, we present the ongoing LoCa approach to context-aware monitoring applications in digital homes. An important constraint in this scenario is that users (patients) are mobile, which means their context might frequently change. Therefore, the way data — coming from different soft- or hardware sensors — is analyzed needs to be automatically adapted, if necessary. The same is true for the interaction of the user with the system. The basis of these adaptations is a powerful context model and its exploitation to dynamically adapt i.) user interfaces and services and ii.) process-based distributed applications (workflows).

The remainder of this paper is organized as follows: Section 2 introduces the LoCa context model. The architecture of the LoCa system is presented in Section 3. In Section 4, we discuss context-aware adaptation in LoCa. The status of the current implementation is presented in Section 5. Section 6 surveys related work and Section 7 concludes.

2. Context Model

LoCa exploits a generic context model to improve health care applications and to facilitate the treatment of patients, both in home care and in stationary care. To reach this goal, we need to adapt processes and user interfaces automatically according to the current context. This, in turn, necessitates the proper representation of context information. We have designed a generic context model for context data management. Figure 1 depicts this model in Entity-Relationship notation. In here, we closely follow the well established definition of context by Day et al. [1]: *Context is any information that can be used to characterize the situation of a subject. A subject is a person, place, or object that is considered relevant to the interaction between a user and an application [...].*

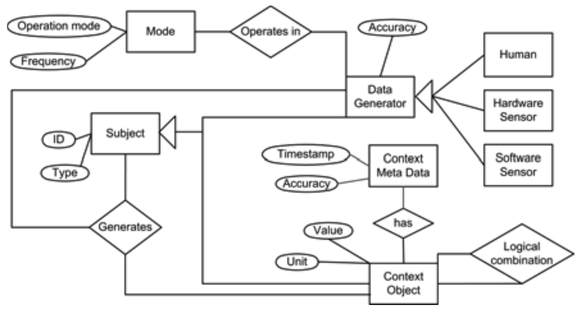


Figure 1. LoCa Context Model

The *Subject* can be a patient, a mobile phone, or an ECG sensor. Conversely, profile data, the medical history, current ECG data, or the current location are examples for context information about a patient. The entity *Context Object* represents the actual context data, e.g., the value of the current location, a document of the medical history, and so on. In order to support data analysis, we store optional meta data about context objects, such as time stamps and data accuracy (which usually depends on the type of sensor used).

The entity *Data Generator* (humans, hardware sensors, software sensors) is designed to capture data about the instrument (sensor) which produces context data: a data generator generates context data about subjects. While many data generators generate atomic data, some sensors may produce compound context objects. For instance, the (GPS) location usually consists of multiple values, such as longitude, latitude, altitude, speed, and bearing. Furthermore, software sensors can combine different kinds of context objects to compose higher level context data. An alarm in case of cardiac problems could be combined of information about the current activity of a patient and his current ECG values. This is covered in the model by means of the relationship *logical combination*.

The context model is able to handle different kinds of context objects, including nested context objects. An important feature of the context model is its rather simple, yet expressive structure. It is powerful enough to cover all the different context objects that have been identified in the requirements analysis phase of LoCa in which several home care and stationary care use cases have been analyzed together with stakeholders from the eHealth domain. Nevertheless, the model can be extended by adding new data generators and thus also new context objects, if necessary.

3. Architecture of the LoCa Platform

Context awareness requires that the information gathered from distributed sensors is stored in a global, albeit distributed database on the basis of the schema presented in Sec. 2. Prior to inserting raw sensor data into this database, it needs to be cleaned and transformed into the global schema. Since context data is a vital input for all LoCa applications, the context data management layer forms the basis of the LoCa architecture depicted in Figure 2.

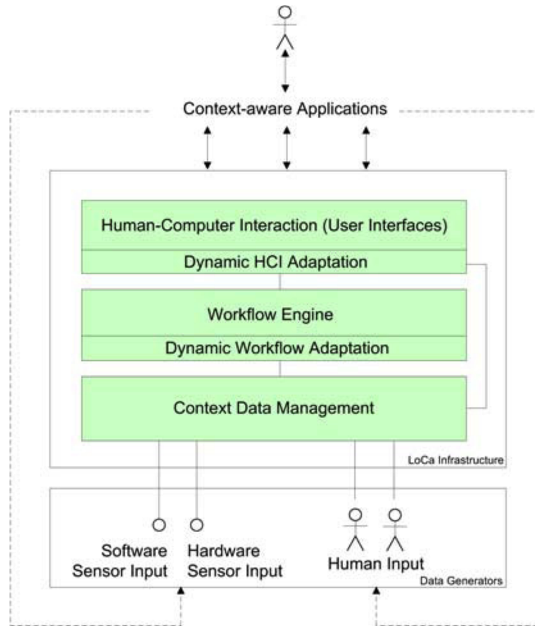


Figure 2. LoCa Conceptual Architecture

On top of context management, the LoCa applications are defined as workflows. The basic assumption is that functionality is available in the form of (web) services so that workflows can be defined by combining existing services. Since complete workflows again have a service interface, service composition can be applied recursively. A crucial part of this layer is dynamic workflow adaptation. This layer makes use of the raw sensor data and their relationships stored in the context layer. The top-most layer of the LoCa architecture deals with the dynamic generation and adaptation of user interfaces. Again, this layer directly accesses the underlying context data management.

All layers are embedded in the LoCa infrastructure which is described in more detail in Section 5. The LoCa architecture offers a unified interface for (individual, user-defined or pre-existing) workflow based applications. According to the context model, LoCa workflow-based applications themselves can be considered software sensors, i.e., they might produce context objects which are subsequently needed for dynamic adaptation.

4. Context-aware Adaptation in LoCa

In what follows, we address the dynamic adaptation needed in LoCa for applications in the eHealth domain, namely at workflow (process) and at user interface level.

4.1. Context-aware Workflows

Traditional approaches to workflow management usually consider static settings as they can be found in business processes or office automation. However, these approaches are far too rigid to handle highly dynamic environments as they can occur in the medical domain, especially when monitoring mobile patients in their (digital) home environment. From a workflow management perspective, these applications are characterized by a potentially large number of i.) exceptions or unforeseen events (e.g., abnormal deviations in sensed physiological data that may require alternative medication); ii.) different ways to achieve a goal (e.g., different devices can be used to meter blood pressure); iii.) decisions only decidable at run-time (e.g., results of tests cause different subsequent tests or treatments); and iv.) dynamic and continuous changes (e.g., new devices, or treatment methods).

Context-aware, adaptable workflows offer much more flexibility than traditional workflows as they allow for structural changes based on evolving user context. Basically, structural changes of workflows can be done at build-time (prior to the instantiation of workflow processes) and at run-time (changing an instance of a workflow). Build-time changes cover evolutionary changes of processes but also changes caused by context changes like new methods of treatments, hospital guidelines, laws, etc. These kinds of workflow changes are not in the primary focus of LoCa. We will mainly address run-time changes such as, for instance, allergic hypersensitivity of patients that cause changes in the treatment process (e.g., adding an allergy test).

There are two kinds of run-time changes [19] — process adaptation and built-in flexibility. Process adaptation, that can be performed at run-time, is based on modification operations like add, delete, or swap of process fragments. Built-in flexibility supports the exchange of process fragments of a workflow. For instance, assume the examination of a special disease differs depending on the age of the patient because the risk to get this disease and its severity increases with the age of the patient. Thus, the examination always follows the same basic structure while the concrete steps depend on the patient's risk group. Therefore, a workflow consisting of placeholders and concrete steps is defined at build-time. Steps that differ depending on the age are defined as placeholder activities and steps that not differ as usual activities. At run-time, placeholder activities are replaced by the concrete fragments depending on the patient's risk group.

Variants of built-in flexibility are described in [19]. Three of them are of particular importance for the eHealth applications in LoCa: i.) late selection, ii.) late modeling, and iii.) late composition. They differ in the degree of decision deferral and need for user experience. The least flexibility is offered by late selection where workflows, defined at build-time, contain placeholder activities that are substituted by a concrete implementation during run-time. Late modeling additionally supports modeling of placeholder activities at run-time. The most flexible pattern is late composition. At build-time, only process fragments are specified. At run-time workflows are composed out of the process fragments available. In LoCa, we will adopt late composition and will make use of the services' semantics (using semantic web service standards) for the actual selection.

Applied to the scenario presented in Section 1, the treatment workflow has to be adapted dependent on the vital parameters of the patient. Assume that the therapy is less successful than expected so that the physician decides to also meter the blood pressure of the patient. In this case the workflow for controlling the patient's health state has to be extended accordingly. Usually, the physician is informed about irregularities in the patient's ECG values by visually highlighted values and, if severe problems occur, by an SMS to his mobile phone. The extension to a new sensor requires also the adaptation of the signal processing and triggering.

In LoCa, we focus on run-time changes of workflows without manual intervention. Particularly, we will provide rules for automated adaptation of workflows, that is, automated fragment selection or composition based on user context and service semantics.

4.2. Context-aware User Interfaces

Adapting user interfaces in a context-aware environment allows the various actors of the system the best possible utilization of the available resources. Therefore simply defining one standard user interface (UI) design and adapting it to the display of the device the user is currently using will not be sufficient [10, 13, 21].

In LoCa, each user interface component (i.e., button, pulldown menu, picture) will be described in an artifact and be interpreted at run-time. This generic description contains the type of the component, its position within a hierarchy, a mapping to the environment that allows listening to incoming information and a label.

Another artifact with a set of rules is responsible for mapping the generic composite to a concrete representation for a given situation. This rendering mechanism is executed at run-time in order to choose the currently most optimal way to display the component. It takes into account the following contextual information: i.) *device:* information about the current device, such as displaying capabilities, current network bandwidth and latency, CPU usage, remaining battery time, etc. This might be a mobile device of the patient or any device of the patient's digital home environment; ii.) *user:* who is using the current device. This information may also cover several users, such as the doctor and a patient during a ward visit; iii.) *location:* the current location of the device may also influence the rendering of a component; iv.) *reason:* the reason why a component is displayed may be difficult to obtain. Possible elements of such an information may be the current calendar entries or tasks, the current patient situation, such as ECG; v.) *time:* the dimension time is not simply a timestamp, but may also include time spans or semantical information, such as "after lunch" or "night".

For the application scenario presented in Section 1, this means for instance that the patient's mobile device knows, by making use of the calendar stored on it, that a specific process needs to be started. The device displays the input fields required to enter the required physiological parameters. If the input field for the blood oxygen saturation value is able to find a viable hardware sensor in its proximity (oximeter), it automatically reads the value from that device, sets itself immutable and moves to the bottom of the display. The mandatory input fields that cannot be processed automatically must be filled in by the patient. Each input component must also decide how to react if, for example, the patient fills in a value before it could find a matching hardware sensor in its environment.

5. Implementation

The implementation of the LoCa infrastructure is currently ongoing. LoCa will use and further advance the open service-oriented infrastructure OSIRIS[1] NEXT (ON)[2]. Originally based on the hyperdatabase vision [16], many ideas from process management, peer-to-peer networks, database technology, and Grid infrastructures were integrated in the past in order to support distributed and decentralized process management [18]. More recent work aims at i.) support for distributed data stream management [5, 6] and ii.) the integration of semantic technologies to enable new ways for flexible and automated process management support. This includes support for distributed and decentralized execution of processes in dynamic (mobile) environments [11] as well as an advanced method to enable automated forward-oriented failure handling [12].

In the context of the LoCa project we will exploit and extend the process management system that has been integrated into ON. It allows for dynamically distributed and decentralized execution of composite semantic services that are described based on OWL-S. On top of this, the user interface will be built based on the Android platform[3].

ON essentially represents a P2P-based open service infrastructure. At its bottom layer it realizes a message-oriented middleware enabling arbitrary services which are deployed at peers to interact by means message exchange. Besides the possibility for end-to-end interactions, the platform also realizes a publish-subscribe messaging paradigm. Furthermore, it incorporates advanced concepts for eager and lazy data replication, taking into account user specified data freshness properties. The platform provides several built-in system services that are used to manage meta and runtime information about the services offered by the peers in the network [18].

ON is fully implemented in Java. One of its key properties is its a small systems footprint (in particular regarding memory) and its internal design is strictly multithreaded in order to take advantage of multi-core CPU technology. Every service spawns its own thread group. Internal message processing is similar to the SEDA approach [20]. It can be deployed in a stand-alone mode on a wide range of devices, starting from mobile platforms, netbooks, up to enterprise computing machines. Moreover, ON can also be deployed as an agent in the JADE[4] agent platform, thus, enabling FIPA compliant usage.

For evaluation and demonstration of our approach, especially of our use cases, we are building a prototype based on Android cell phones. Figure 3 shows an early prototype of the user interface for a physician.

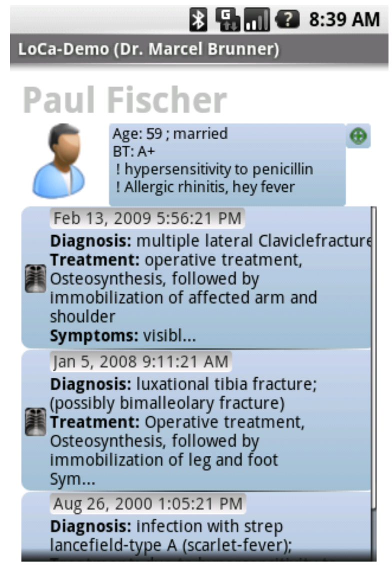


Figure 3. Screenshot of LoCa Demonstrator

6. Related Work

In the last years, a number of projects have been carried out in the eHealth domain. In particular, many projects

[1] Open Service Infrastructure for Reliable & Integrated process Support
[2] http://on.cs.unibas.ch
[3] http://source.android.com
[4] http://jade.tilab.com

apply workflow and process technology for distributed application in eHealth. Akogrimo [8] deals with the support of dynamic virtual organizations that require the ability to change its structure dynamically and to access data from mobile resources. ADEPT [15] allows to dynamically change the type of workflow instances in order to react to changes in the application (e.g., patient's therapy). While ADEPT addresses mainly change patterns, CAWE (Context Aware Workflow System) [3] deals with built-in flexibility.

A number of eHealth projects also take into account context. The MARC project [2] provides a passive monitoring system that can be used for elderly people. CodeBlue [7] explores various wireless applications in the eHealth domain with a focus on 3D location tracking. ARCS [17] addresses user interface adaptation in eHealth applications. It provides web-based interfaces mainly for stationary devices for manual disease monitoring. In [4], eHealth applications and services to support mobile devices have been designed.

Online monitoring and streaming data is more and more emerging in eHealth. The MyHeart project [14] monitors cardio-vascular parameters using measuring wearable devices (i.e., devices that are integrated into clothes). The PHM project [9] measures different vital parameters either continuously or at determined time intervals.

7. Conclusion and Future Work

LoCa is an ongoing effort that will provide a novel approach to context and location-aware eHealth applications as they can be found when monitoring physiological data and activity status of patients in a digital home environment. By providing generic support for the context-aware adaptation of workflows and user interfaces, LoCa is intended to be applied to other scenarios as well, e.g, in stationary care.

In close collaboration with healthcare practitioners and experts from industry, we have identified several concrete scenarios. The requirements coming from there will be considered when completing the implementation of the LoCa system based on the ON platform. Finally, these scenarios will be evaluated together with our medical partners.

References

[1] G. Abowd, A. Dey, P. Brown, et al. Towards a Better Understanding of Context and Context-Awareness. In *Proc. Int'l Symp. on Handheld and Ubiquitous Computing*, pages 304–307, London, UK, 1999. Springer.

[2] M. Alwan, S. Kell, B. Turner, et al. Psychosocial Impact of Passive Health Status Monitoring on Informal Caregivers and Older Adults Living in Independent Senior Housing. In *Proc. ICTTA'06*, pages 808–813, 2006.

[3] L. Ardissono, R. Furnari, A. Goy, et al. A Framework for the Management of Context-aware Workflow Systems. In *Proc. WEBIST 2007*, pages 80–87, 2007.

[4] J. Bardram. Applications of Context-Aware Computing in Hospital Work – Examples and Design Principles. In *Proc. ACM SAC*, pages 1574 – 1579, 2004.

[5] G. Brettlecker and H. Schuldt. The OSIRIS-SE (Stream-Enabled) Infrastructure for Reliable Data Stream Management on Mobile Devices. In *Proc. SIGMOD*, pages 1097–1099, June 2007.

[6] G. Brettlecker, H. Schuldt, and H.-J. Schek. Efficient and Coordinated Checkpointing for Reliable Distributed Data Stream Management. In *Proc. ADBIS'06*, pages 296–312, Thessaloniki, Greece, 2006.

[7] T. Gao, D. Greenspan, M. Welsh, et al. Vital Signs Monitoring and Patient Tracking Over a Wireless Network. In *Proc. IEEE EMBS*, pages 102–105, 2005.

[8] T. Kirkham, D. Mac Randal, J. Gallop, and B. Ritchie. Akogrimo — a Work in Progress on the Delivery of a Next Generation Grid. In *Proc. SOAS'05*, 2005.

[9] C. Kunze, W. Stork, and K. Müller-Glaser. Tele-Monitoring as a Medical Application of Ubiquitous Computing. In *Proc. MoCoMed'03*, pages 115–120, 2003.

[10] P. Langley. User Modeling in Adaptive Interfaces. In *Proc. UM'99*, pages 357–370. Springer, 1999.

[11] T. Möller and H. Schuldt. A Platform to Support Decentralized and Dynamically Distributed P2P Composite OWL-S Service Execution. In *Proc. MW4SOC'07*. ACM, 2007.

[12] T. Möller and H. Schuldt. Control Flow Intervention for Semantic Failure Handling during Composite Serice Execution. In *Proc. ICWS'08*, pages 834–835, 2008.

[13] E. G. Nilsson, J. Floch, S. O. Hallsteinsen, and E. Stav. Model-based user interface adaptation. *Computers & Graphics*, 30(5):692–701, 2006.

[14] M. Pacelli, G. Loriga, N. Taccini, and R. Paradiso. Sensing Fabrics for Monitoring Physiological and Biomechanical Variables: E-textile solutions. In *Proc. Int'l Symposium on Medical Devices and Biosensors*, pages 1–4, 2006.

[15] M. Reichert, S. Rinderle, and P. Dadam. On the Common Support of Workflow Type and Instance Changes Under Correctness Constraints. In *Proc. CoopIS'03*, pages 407–425, Catania, Italy, Nov. 2003. Springer LNCS.

[16] H.-J. Schek and H. Schuldt. The Hyperdatabase Project - From the Vision to Realizations. In *Proc. BNCOD*, pages 207–226, Cardiff, UK, 2008.

[17] G. Schreier, A. Kollmann, M. Kramer, et al. *Computers Helping People with Special Needs*, volume 3118, pages 29–36. Springer, 2004.

[18] C. Schuler, C. Türker, H.-J. Schek, R. Weber, and H. Schuldt. Scalable Peer-to-Peer Process Management. *Int. J. of Business Process Integration & Management*, 2006.

[19] B. Weber, S. Rinderle, and M. Reichert. Change Patterns and Change Support Features in Process-Aware Information Systems. In *Proc. CAiSE 2007*, pages 574–588. Springer LNCS, Trondheim, Norway, June 2007.

[20] M. Welsh, D. Culler, and E. Brewer. SEDA: An Architecture for Well-Conditioned, Scalable Internet Services. In *18th Symp. on OS Principles*, Banff, Canada, 2001.

[21] Z. Yu, X. Zhou, D. Zhang, et al. Supporting Context-Aware Media Recommendations for Smart Phones. *IEEE Pervasive Computing*, 5(3):68–75, 2006.

An Intelligent Web-based System for Mental Disorder Treatment by Using Biofeedback Analysis

Bai-En Shie[1], Fong-Lin Jang[2], Richard Weng[3], Vincent S Tseng[1,4]*

[1]Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan, Taiwan, R.O.C..

[2]Department of Psychiatry, Chi-Mei Medical Center, Tainan, Taiwan, R.O.C..

[3]Innovative DigiTech-Enabled Applications & Services Institute, Institute for Information Industry, Kaoshiung, Taiwan, R.O.C..

[4]Institute of Medical Informatics, National Cheng Kung University, Tainan, Taiwan, R.O.C..

*Correspondence E-mail: tsengsm@mail.ncku.edu.tw

Abstract—**With the rapid development of the communication technology, the Internet plays a more and more important role in many applications of healthcare. In healthcare field, mental disorder treatment is an important topic, with cognitive behavioral therapy and biofeedback therapy as two emerging and noteworthy methods. A number of researches on the integrations of the Internet and mental healthcare have been proposed recently. Thus, this research aims at the development of an online treatment system for panic patients by combining the cognitive behavioral therapy, biofeedback therapy, and web technologies. The system provides a more convenient communication between patients and medical personnel. The essential treatments and some related information provided by the medical personnel can be downloaded or used online by the patients via a web-based interface. On the other hand, important information such as physiological data can also be uploaded to the server databases automatically. Therefore, considerable time on the treatments can be saved for both patients and therapists, and the medical costs can be highly reduced. The experimental results show that the curative effects of the mental disorder patients are highly depend on the physiological status. The results of this research are expected to provide useful insights for the field of mental disorder treatment.**

Keywords—**biofeedback analysis, online mental therapy, mental disorder treatment, intelligent healthcare, data mining**

I. Introduction

Since the rapid development of the electronic communication technology, the Internet plays a more and more important rule in the domain of medicine, especially in the healthcare field. A number of researches on the integration of the Internet and the treatments for mental healthcare were proposed [1, 2, 3, 4, 5, 6, 8, 14, 15, 16]. In these researches, the treatments with the Internet were mostly used for melancholia and anxiety disorders. In addition, there were also some researches that applied the Internet for the treatments of the patients with the substance use disorders, such as smoking [4, 6] and alcoholism [5, 14]. That is, the researchers applied the Internet to cognitive behavioral therapies. This can not only save the time of the patients and therapists while achieving the goal of treatments, but also reduce the cost of healthcare.

Recently, various mental disorders become more and more popular in modern societies. Mental disorders are mainly divided into major and minor disorders. Minor mental disorders are mainly expressed in affective disorder, such as anxiety, depression, and thought disorder, such as obsession. However, the patients' cognitive thinking, the ability of logic inference and self-checking ability are generally normal. The patients of major mental disorders may show anxiety and obsession in the initial stage, but their cognition will be very bad with the self-checking ability almost lost. The common minor mental disorders are such as anxiety disorder, obsessive-compulsive disorder, depression, and phobia, and on the other hand, the common major mental disorder is such as schizophrenia.

In mental disorders, panic disorder is a kind of chronic disease. It is a common disease of the cases in the hospital emergency-rooms. The mental symptoms of the patients with panic disorders are fear of losing control of themselves, derealization, depersonalization, and the feeling of impending death. The physiological symptoms of the patients are dizziness, tyspnea, tachypnea, and palpitations. The patients will be very fearful and uncomfortable. Some severe patients will even afraid of going out to avoid appearing outdoors such as in open space, on bridges, in queues, in cars, in crowd, or other places which are difficult to escape from people [17]. For the patients with panic disorders, the symptoms will be repeated constantly and unexpectedly, which make the sufferer feel highly distressed and apprehensive. Therefore, their behaviors will be blatantly obviously. They will endeavor the occasions which they afraid. In the later stage, they will even be melancholy and agoraphobic and these may result in decreasing of their family functionalities. The symptoms of panic disorders are not easily diagnosed. They are often diagnosed as heart attack or other diseases and the patients may have many unnecessary medical check-ups. These symptoms not only waste medical resources and delay time limitations for treatments, but also results in inconveniencing of social and occupational functionalities of the patients [10, 20].

In view of these, we aimed at building an intelligent mental disorder treatment system with the integration of cognitive behavioral therapy, biofeedback therapy and web technologies in this paper. The main contributions of this paper are as follows. First, the system provides a convenient interface for the communication between patients and hospital staffs. Second, the hospital staffs enable patients to query or download information via the Internet. Third, the patients can upload their physiological data and self-rating scales to the databases of the hospitals via the Internet.

For biofeedback measurements, we used a new biofeedback device, named *emotion ring*, as shown in Figure 1 to record the patient's finger skin temperature. Different to other biofeedback devices, the advantages of the emotion ring are compact size, easy to carry, ease to operation, and wireless data communication. We applied online progressive muscle relaxation training combined with the emotion ring measuring to help patients learn how to relax themselves and alleviate the symptoms of panic disorder. Once the patients learn the somatic cues for relaxation and the method to obtain rapid relaxation, they were able to apply the methods and cues to relieve the symptoms of panic disorder. Moreover, we used the proposed online therapy system for the patients to perform the treatment courses themselves at home. We also requested them to upload the biofeedback data via the system daily and whereby therapists could quickly manage patients' latest data. Furthermore, patients were asked to fill out the self-rating scales online and upload them for the therapists, so that the therapists could know the patients' mental status, judge their curative effect, and give them some necessary feedbacks.

Fig. 1. The biofeedback device: emotion ring.

This paper is the first research for the system of integration of cognitive behavioral therapy, biofeedback therapy and the Internet. We expect the system can be used by the patients to practice biofeedback therapy at home. In this paper, we also constructed a complete biofeedback online therapy model, which was composed of cognitive behavioral therapy, data transmission and storage, and connecting and interacting between patients and therapists via the Internet. The results are expected to increase the convenience of mental therapy, decrease the medical cost, be able to deal with more patients who need mental therapy and provide a beneficial application for public health in society and also academia. In the experiments section, we employ the data to explore the possibility of giving mental healthcare with physiological data. We expect that the system can assist the prevention and treatment of mental disorders by monitoring the physiological data with real clinical verification.

The rest of this paper is organized as follows. In Section 2, we summarize the existing researches on panic disorders. In section 3, we describe the proposed online treatment system for panic disorders in detail. The performance study of our research is presented in Section 4. Section 5 is the conclusion of the paper.

II. Related Works

Panic disorder is encountered frequently in general medical practices and emergency services. The data from National Comorbidity Survey Replication of the United States showed the lifetime prevalence estimates are 3.7% of panic disorder without agoraphobia (panic disorder only), and 1.1% of panic disorder with agoraphobia [7]. The international lifetime prevalence rates of panic disorder ranged from 0.13% in rural village of Taiwan to 3.8% in the Netherlands [18]. This disorder is rather debilitating to the sufferer, and even causes depression or suicide [20]. The life quality of the victims of panic disorder is dismal, and even worse than those with major depression [10]. The victims of panic disorder also received more welfare or some form of disability compensation [13].

For public health, the optimal treatment for panic disorder is an important task to be dealt with. In clinical practice, two major modalities have been applied to its treatment of panic disorder: one is pharmacotherapy and the other is non-pharmacological psychotherapy. For psychotherapy, cognitive behavioral therapy is the main mode and has been proved to be effective for symptom management and prevention of recurrence for panic disorder [17, 21]. Thanks to the advancement in computing and the Internet, computer-aided cognitive behavioral therapy has been employed for more than one decade. It is any computing system that aids cognitive behavioral therapy to make computations and treatment decisions [11]. But computer-aided cognitive behavioral therapy should not only expedite communication or overcome the problem of distance; it consists of computation rather than replacing routine paper leaflets only [12].

Most Internet interventions for mental disorders are cognitive behavioral programs that are proposed as guided self-help programs on the Internet. Randomized controlled studies on the use of Internet interventions for the treatment of mental disorders are still scarce [15]. From the limited literature it showed that computer/Internet-aided cognitive behavioral therapy was superior to waiting lists and placeboes assignment across outcome measures, and the effects of computer/Internet-aided cognitive behavioral therapy were equal to therapist-delivered treatment across anxiety disorders. However, conclusions were limited by small sample sizes, the rare use of placebo controls, and other methodological problems [16].

Treating panic disorder sufferers via the Internet is a rational concept, not only considering the issue of transportation of patients but also that of those suffering from agoraphobia. Up to date, publications about clinical trials of Internet-based cognitive behavioral therapy for panic disorder were mainly from Sweden, United Kingdom, and Australia. Carlbring et.al constructed a cognitive behavioral therapy treatment program consisting of stepwise intervention modules: psychoeducation, breathing retraining and hyperventilation test, cognitive restructuring, interceptive exposure, exposure in vivo, and relapse prevention [1]. The participants got significant improvement in all dimensions of measures. They further compared an Internet-based treatment program with an applied relaxation program which instructed the participants on how to relax expediently and applying relaxation techniques to prevent a relapse into a panic attack [2]. The applied relaxation condition has a better overall effect compared to the cognitive behavioral therapy program, and the effectiveness of the two groups was similar. Recent randomized trials demonstrated that Internet-based cognitive behavioral therapy for panic disorder could be as cogent as traditional individual cognitive behavior therapy [3, 8].

III. Proposed Methods

In this paper, we integrate our mental disorder therapy system with the Internet to efficiently collect the biosignal data, the self-rating scales and the personal profiles of the patients with mental disorders. In this section, we describe the scenario and the functions of our proposed online therapy system.

A. *User Scenarios*

There are four kinds of users in this system: patients, therapists, hospital managers, and system managers. In the following, we explain the user scenarios in detail.

1) The scenario of the patients with mental disorders: The patients with mental disorders use the finger temperature measurement system and upload the results to the databases daily. Either weekly or monthly, they need to fill out the self-rating scales which are provided by the therapists in the system. The patients can also query their own treatment records or see the suggestions which were provided by their therapists.

2) The scenario of the therapists: The therapists use the system to manage the data uploaded by patients. The data are composed of the finger temperature which is measured by the patients daily and the self-rating scales which are filled out by the patients weekly or monthly. The therapists can also reply some suggestions to the patients after observing the data. When the patients afterward login the system, they can check the suggestions via the system conveniently. Besides, the therapists can create new accounts of patients by themselves without operations by the database managers. When a patient finishes the treatment procedure, the therapist can directly close this case in the online system.

3) The scenario of the hospital managers: If a hospital manager is also a therapist, he/she can manage his/her patients like a therapist does in the system. Besides, the hospital managers can also manage all therapists in each hospital via the system. The hospital managers can create new therapists' accounts by themselves without contacting the database managers.

4) The scenario of the system managers: The system managers do not actually need to use the system. They just manage and maintain the system. They can create new accounts for hospital managers. However, since the treatment records can not be made arbitrarily public, the system managers can not see patients' treatment data.

The login roles of the users are shown in Figure 2. In the figure, we can know the top management of this system is system manager. He/She can create the account of the hospital managers. For each individual hospital, there is only one hospital manager handling all the therapists who use the system in the hospital. By the way, the therapists can manage all their own patients via the system.

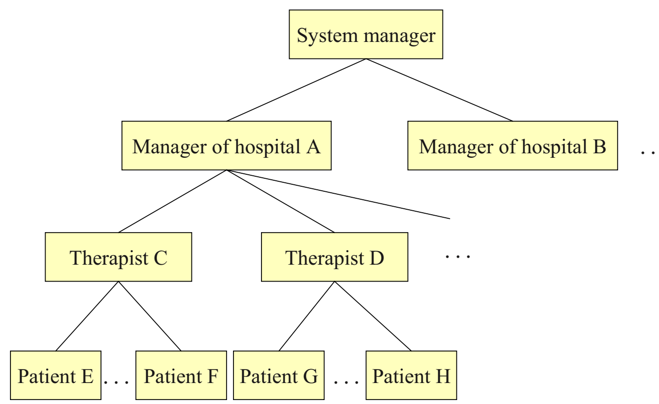


Fig. 2. Sketch map of login roles.

B. *The techniques about measuring finger temperature*

In the following, we describe the communication processes and methods between the emotion rings and the computers. First, we install the device driver of the emotion ring. After installation, the MAC address of the emotion ring and the detected temperature will be transmitted from the emotion ring to the USB receiver once a second. When the USB receiver gets data, it simulates a COM port and transmits the data with 11 bytes. Table I is an example of the transmitted data. The first byte is fixed as "A3". The second to the ninth bytes are the MAC address of the emotion ring. The last two bytes are the temperature data. The first four MAC address of all emotion rings are all the same, "001CD902". The received temperature data are ten times of the actual temperature.

TABLE I
Example of the Data Transferred by the Emotion Ring.

Preamble	MAC Address								Data	
A3	00	1C	D9	02	00	00	00	3B	01	0A
1 Byte	8 Bytes								2 Bytes	

The execution environment of the receiving end is Java applet. The basic libraries of Java do not support the input and output of the serial ports. User's Java environments will be detected and the libraries are created. The program for the receiver needs to search a free COM port for receiving data. After receiving data, the received information from the last eleven bytes to the last seven bytes are checked instead of from the first to the fifth bytes. The reason for this being if some errors occur during data transmission, the receiver may receive the data from the middle of the previous data instead of the first byte of the latest data. So we check from the last of the received data to avoid any error occurrence.

After checking the received data, the program acquires the data from COM port once a second, and output the number which is one-tenth of the last two bytes of data. Table I is an example of received data. The decimal in the last two bytes of the data, i.e., 010A, is 266, and its one-tenth is 26.6. This indicates the temperature which is detected at that time is 26.6℃. However, sometimes the USB receiver may not be given the data due to poor signaling strength. The emotion ring will be regarded as "not exist" when the program does not detect any data after three seconds.

C. System Workflow

First, users enter and login to the system website. Figure 3 is the screenshot of the patients' homepage. In the webpage, we remind the patient whether or not he/she has completed the daily course. If the patient does not complete it, the instructions in the related webpages will lead he/she to do so. If there are some self-rating scales to be completed, it will also be mentioned in the homepage. In this way, the patients will not forget the routine task they need to complete on that day. If the patients want to query their previous finger temperature results records or self-rating scales, or view therapists' suggestion, they can find them on the "records review" pages. On the page "contact therapist", the contact information for the therapists, such as e-mails, is provided for the patients.

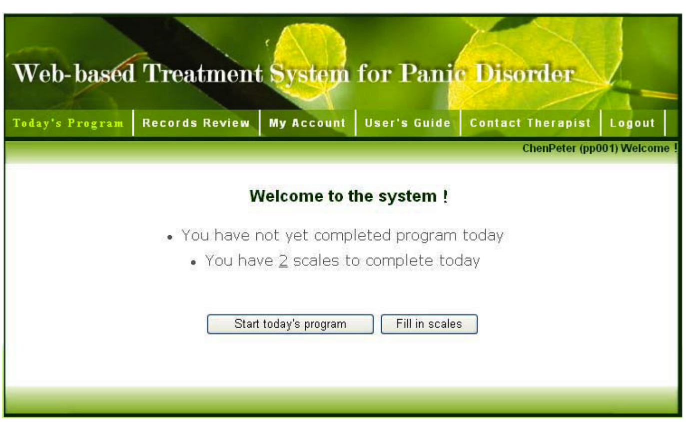


Fig. 3. Screenshot of the homepage for patients.

The system flowchart is shown in Figure 4. As can be seen the patients can select the functions arbitrarily when they login to the system. For convenience in using the system and decreasing confusion for the users, the system details the options and procedures the user has to complete on that day. Through the guidance of the system, a patient may use the system as follows: First, he logs into the system and is informed by the homepage that he hasn't completed the daily treatment course on that day. Then he completes it and uploads the temperature data to the system database. In the next, he returns to the homepage and finds that he has a self-rating scale to complete, so he completes it. After finishing that day's necessary tasks, he goes to the pages to see the suggestions which his therapist has given the previous day, the results of finger temperature and the self-rating scales are then uploaded that day. In the end he logs out of the system.

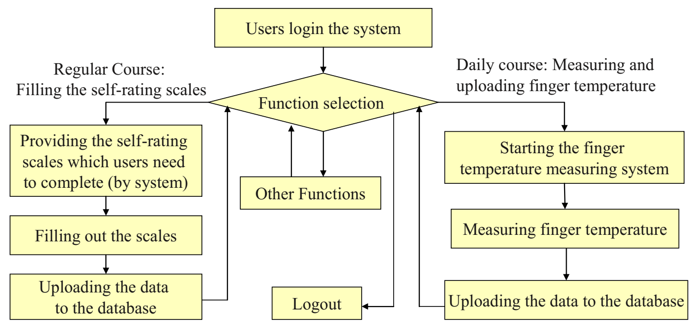


Fig. 4. Flowchart of the system (for patients).

As shown in Figure 4, the main functions of the system are measuring finger temperature, filling out self-rating scales, and uploading the data. The function of measuring temperature is integrated into the online therapy system. The patients can just click the "start measurement", "pause measurement" or "end measurement" buttons, and can then easily complete the required tasks respectively. After the measurements, the data is uploaded to the database automatically by the system. This avoids any kind of confusion for the users from other various unconnected types of programs, such as one for measuring temperature and another for uploading data. The simplicity of the system promotes a willingness by the patients to participate in this system, which in turn popularizes it will the participating patients.

On the other hands, the hospital manager may also be a therapist, so some user functions of the hospital manager and the therapist are the same. The main functions of therapists are managing the patients, which entails viewing the data daily, replying to suggestions from the patients, viewing their periodical self-rating scales, filling out the patients' self-rating scales, adding new cases, and so on. Besides the above functions, the main functions of the hospital managers are adding new therapists and managing of them.

IV. Experimental Evaluation

In this section, we introduce the sources, the designs, the results and the discussions about the research.

A. *The Real Data for Experiments*

In the experimental analyses, we use the data obtained from subjects from the department of psychiatry in a medical center in Taiwan. In this research, we gave each patient a muscle relaxation course, i.e., muscle relaxation music, a biofeedback device, i.e., the emotion ring, and an account for login into the system. The patients were asked to practice the online treatment courses and upload the daily results every day. The patients would also upload the scores of their emotions before and after the courses and also the feelings during the courses to the database. The therapists would review the data periodically and give the patients some feedback or suggestions if necessary.

In this research, the patients were divided into an experimental group and a control group. The patients in the experimental group did the courses as mentioned above, i.e., listening to muscle relaxation music and in the meanwhile measuring the finger temperature. On the other hand, the patients in the control group just listened to the muscle relaxation music without temperature measuring. The control group was mainly used for verification in the experiments.

During this research, we collected the patients' personal profiles and physiological data by different mechanisms. Among them, the physiological data were extracted and collected by the emotion rings. After data collection, we utilized our data mining system for analyzing the data. Before the analyses, we did the preprocessing on the collected data. At this step, we focused on the missing data and processed essential data cleaning and some integration on them. For example, some data would be stored by another format or the redundant and missing data would be deleted. Thus, the processing time is reduced and the accuracy of experiments is enhanced.

B. *Experimental Design*

In this part, we describe the data analysis method for the collected data, i.e., the patients' profiles and the biofeedback data. We integrate the data mining techniques with the processional knowledge of the mental disorder to design methods of the data mining analysis.

The proposed data mining analysis is the association analysis of curative effect and the biofeedback data. The framework of this analysis is shown in Figure 5. We analyze the association between the biofeedback data extracted from the emotion rings and the curative effects. In this analysis, the finger temperature data is regarded as time series data. We apply the SAX algorithm [9] to transform the numerical data to sequence data. After data transformation, we apply sequential pattern mining to the sequence data for finding sequential patterns. Then we apply the CBS algorithm [19] for building classification models on curative effects. The results could be useful references in assisting the therapists in predicting the curative effect by the treatment conditions.

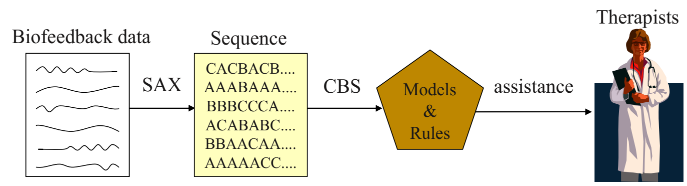


Fig. 5. The framework of the analysis of the curative effect and biofeedback data.

C. *Experimental Results*

In this part, we address the experiment results of the analysis of the curative effect and the biofeedback data. We use real datasets as mentioned above for the analysis. Before the analysis, we apply data preprocessing methods to prune the missing or error data as follows. For a tuple whose temperature differs from the previous one by more than 2℃, it will be considered as an error and then pruned. Naturally, the temperature difference of a human will not be above 2℃ in one second. This happens in the data because the battery of the device is flat or the patients interrupt the course, such as the emoting ring is suddenly removed from the finger.

With regards to the curative effects, we use two types of scores for objectively and subjectively judging them. One is self-rating scores which are determined by patients themselves, and another is the curative effects which are determined by the patients' therapists. We perform the following two experiments by different conditions as follows.

Experiment A. In this experiment, we take all patients' biofeedback data. We set the class for each tuple according to the patients' self-rating scores. If the scores after the courses are better than the scores before, we regard the treatment effects as "good"; otherwise, they are considered as "bad." The class values of the tuples in this experiment are just good or bad. After data preprocessing, we divide the data into training data and testing data with the ratio of 7:3. The experimental results are shown in Table II. In Table II and Table III, the column "inner testing" means the accuracy of the training data and "outer testing" means the accuracy of testing data. By Table II, we can see the overall accuracy is high, i.e., above 80%. It can be seen from this that the curative effects are highly dependent on the biofeedback data. Furthermore, we can also know that the biofeedback data can really reflect the patients' mental state. The results could be important for the therapists' diagnosis.

TABLE II
THE RESULTS OF EXPERIMENT A.

	Inner testing	Outer testing
Accuracy	0.83	0.85
Precision of good	0.85	0.92
Recall of good	0.97	0.92
F-measure of good	0.91	0.92

Experiment B. In this experiment, we take all patients' biofeedback data. We set the class to each tuple according to the curative effect which is determined by the therapists. There are three kinds of curative effect which is judges by therapists:

good, bad, and medium. In this experiment, we use the tuples with the class good and medium. We also divide the data into training data and testing data by 7:3. The experimental results are shown in Table III. By Table III, we can observe that the results are a little worse than Experiment A. These is because the therapists took into account not only the patients' biofeedback data and the self-rating scores, but also the patients' feelings and moods during the courses. These might cause some variants on the previous experimental results whose curative effects are judged by using only patients' biofeedback data.

TABLE III
THE RESULTS OF EXPERIMENT B.

	Inner testing	Outer testing
Accuracy	0.81	0.70
Precision of good	0.90	0.86
Recall of good	0.86	0.92
F-measure of good	0.88	0.80

From the above experiments, we can ascertain that the curative effects are highly dependent on the biofeedback data, i.e., the curves of finger temperature, for the patients of panic disorder. By using this system, we can better control the patients' status when they are performing the biofeedback therapies. In other words, we can know not only the patients' physical state but also their mental state when they are participating in the courses.

V. CONCLUSIONS AND FUTURE WORKS

In this paper, we have proposed a web-based online therapeutic system for mental disorders. The contributions of our system are as follows. First, the patients can get the information or services which are provided by the system. Second, the patients can measure and upload their physiological status via the system. Third, the therapists and the hospital managers can manage their patients conveniently via the system. By the experimental results, we can know that the biofeedback data is useful for judging the curative effects of the patients with panic disorders. For the future work, we will apply the system to the mobile platforms such as mobile phones and PDAs so that the users may use this system more conveniently and ubiquitously.

ACKNOWLEDGEMENT

This research was supported by the Applied Information Services Development & Integration project, Phase II of Institute for Information Industry and sponsored by MOEA, R.O.C..

REFERENCES

[1] P. Carlbring, B. E. Westling, P. Ljungstrand, L. Ekselius, G. Andersson. “Treatment of Panic Disorder via the Internet: A Randomized Trial of a Self-Help Program.” Behavior Therapy 2001; 32(4): 751-764.

[2] P. Carlbring, L. Ekselius, G. Andersson. “Treatment of panic disorder via the Internet: a randomized trial of CBT vs. applied relaxation.” Journal of Behavior Therapy and Experimental Psychiatry 2003; 34: 129–140.

[3] P. Carlbring, E. Nilsson-Ihrfelt, J. Waara, C. Kollenstam, M. Buhrman, V. Kaldo, M. SÖderberg, L. Ekselius, G. Andersson. “Treatment of panic disorder: live therapy vs. self-help via the Internet.” Behaviour Research and Therapy 2005; 43(10): 1321–1333.

[4] N. K. Cobb, A. L. Graham, B. C. Bock, et al, “Initial evaluation of a real-world Internet smoking cessation system.” Nicotine Tob. Res. 2005; 7: 207-216.

[5] J. A. Cunningham, K. Humphreys, A. Koski-Jannes, J. Cordingley, “Internet and paper self-help materials for problem drinking: is there an additive effect?” Addict Behav. 2005; 30: 1517-1523.

[6] J. F. Etter, “Comparing the efficacy of two Internet-based, computer-tailored smoking cessation programs: a randomized trial.” J. Med. Internet Res. 2005; 7: e2.

[7] R. C. Kessler, W. T. Chiu, R. Jin, A. M. Ruscio, K. Shear, E. E. Walters. “The epidemiology of panic attacks, panic disorder, and agoraphobia in the National Comorbidity Survey Replication.” Archives of General Psychiatry 2006; 63(4): 415-424.

[8] L. A. Kiropoulos, B. Klein, D. W. Austin, K. Gilson, C. Pier, J. Mitchell, L. Ciechomski. “Is Internet-based CBT for panic disorder and agoraphobia as effective as face-to-face CBT?” Journal of Anxiety Disorders. 2008; 22(8): 1273-1284.

[9] J. Lin, E. Keogh, S. Lonardi, B. Chiu. “A symbolic representation of time series, with implications for streaming algorithms.” 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery, 200325.

[10] J. S. Markowitz, M. M. Weissman, R. Ouellette, J. D. Lish, G. L. Klerman. “Quality of Life in Panic Disorder.” Archives of General Psychiatry 1989; 46(11): 984-992.

[11] I. M. Marks, S. C. Shaw, R. Parkin, “Computer-aided treatments of mental health problems.” Clinical Psychology: Science and Practice 1998; 5: 151-170.

[12] I. M. Marks, K. Cavanagh, L. Gega. “Computer-Aided Psychotherapy: Revolution or Bubble?” The British Journal of Psychiatry 2007; 191(6): 471 - 473.

[13] M. V. Mendlowicz, M. B. Stein. “Quality of Life in Individuals with Anxiety Disorders.” The American Journal of Psychiatry 2000; 157(5): 669-682.

[14] M. J. Moore, J. Soderquis, C. Werch, “Feasibility and efficacy of a binge drinking prevention intervention for college students delivered via the Internet versus postal mail.” J. Am. Coll. Health 2005; 54: 38-44.

[15] C. B. Pull. “Self-help Internet interventions for mental disorders.” Current Opinion in Psychiatry 2006; 19: 50–53.

[16] M. A. Reger, G. A. Gahm. “A meta-analysis of the effects of Internet- and computer-based cognitive-behavioral treatments for anxiety.” Journal of Clinical Psychology 2009; 65(1): 53-75.

[17] P. P. Roy-Byrne, M. G. Craske, M. B. Stein. “Panic disorder.” The Lancet 2006; 368(9540): 1023-1032.

[18] J. M. Somers, E. M. Goldner, P. Waraich, L. Hsu. “Prevalence and Incidence Studies of Anxiety Disorders: A Systematic Review of the Literature.” Can J Psychiatry 2006; 51: 100–113.

[19] V. S. Tseng and C. H. Lee. “CBS: A New Classification Method by using Sequential Patterns.” In Proc. SIAM Int’l Conf. on Data Mining, USA, April, 2005.

[20] M. M. Weissman, G. L. Klerman, J. S. Markowitz, R. Ouellette. “Suicidal ideation and suicide attempts in panic disorder and attacks.” The New England of Journal of Medicine 1989; 321(18): 1209-1214.

[21] J. L. Wetherell, E. J. Lenze, M. A. Stanley. “Evidence-based treatment of geriatric anxiety disorders.” Psychiatric Clinics of North America 2005; 28(4): 871-896.

Adaptive SmartMote in Wireless Ad-Hoc Sensor Network

Sheng-Tzong Cheng[1], Yao-Dong Zou[1], Ju-Hsien Chou[1], Jiashing Shih[1], Mingzoo Wu[2]

[1]Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan, Taiwan
[2]Innovative DigiTech-Enabled Applications & Services Institute, Institute for Information Industry, Kaohsiung, Taiwan

Abstract—This paper describes an update mechanism for large wireless ad-hoc sensor networks (WASNs). In wireless sensor networks, the nodes may have to be reprogrammed, especially for design-implement-test iterations. Manually reprogramming is a very cumbersome work, and may be infeasible if nodes of the network are unreachable. Therefore, a wireless update mechanism is needed. Exchanging the running application on a node by transmitting the complete program image is not efficient for small changes in the code. It consumes a lot of bandwidth and time. The proposed framework, Adaptive SmartMote, defines and supports control JOBs that allow computation, behaviors. The goal of this paper is to use programmable packet to update sensor behaviors. To reduce the code transferred and power consumption, we propose a group management architecture. This architecture helps reduce power consumption and increase node number that control by Leader Node in WASNs. The proposed update protocol has been implemented on the Tmote-based Octopus II sensor node, which is named SmartMote, which runs TinyOS [1], a component-based operating system for highly constraint embedded platform.

1. Introduction

In our daily life, we encounter sensors of all different kinds without even taking notice of. Motion sensors turn on lights when we walk by, the heating or air conditioning of rooms is controlled by temperature sensors and fire detectors alert us in case of emergency.

Recently, a lot of attention has been directed toward extended, “Active” or “intelligent” sensors, that can not only conduct certain measurements, but are equipped with computational power and over-the-air communication. A lot of additional application areas have appeared for these new devices, ranging from medical applications, home automation, traffic control and monitoring of eco-systems to security and surveillance applications.

Researchers have mostly been concerned with exploring applications scenarios, investigating new routing and access control protocols, proposing new energy-saving algorithmic techniques, and developing hardware prototypes of sensor nodes.

Sensor nodes may need to be reprogrammed, e.g. update the running program. An additional module may have to be added to the program, or a complete protocol implementation exchanged. Another important reason is that applications go through a number of design-implement-test iterations during the development cycle. It is highly impractical to physically reach all nodes in a network and manually reprogram them by attaching the node to a laptop or PDA, especially for a large number of distributed sensors. It may also be simply infeasible in various scenarios, if the nodes are located in areas that are unreachable.

A wireless updating scheme is required to set all nodes up to date with the new version of the application. Another consideration is the amount of code transferred. While it is normal to send the whole code if the application needs to be replaced, it does not make much sense in other cases. If we just add or exchange a part of the code, we transmit code that is already available on the node, maybe just shifted from its original location in program memory by a certain offset. Also if a bug has been identified and fixed in the test process, the biggest part of the code remains exactly the same, probably only differing for some functions or constants. To reduce this redundancy, it is much more efficient in terms of used bandwidth and time to only send the changes in the code, and leave the recombination of the new code to the node itself.

2. Related Work

We overview existing approaches vary from single-hop reprogramming over multi-hop reprogramming to complete virtual machine [6] approaches.

2.1 XNP

One of the very first approaches used to reprogram sensor nodes was included in TinyOS. With XNP [2][3][7], mica2 and mica2dot nodes can be reprogrammed over the air. Only complete images can be transferred to the node, since XNP does not consider identical code parts. There is no forwarding mechanism in the program, so only the nodes in the immediate neighborhood of the basis station can be reprogrammed. This is also called single-hop reprogramming.

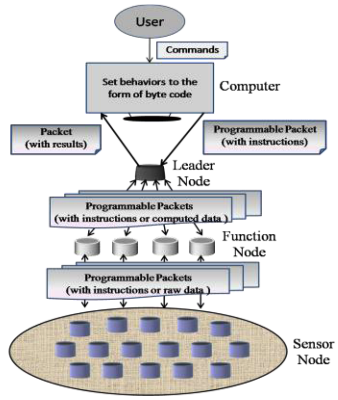


Fig. 1: Adaptive SmartMote packet transmission

2.2 Multi-Hop Over-The-Air-Programming

Multi-hop Over-The-Air-Programming (MOAP) [3] uses basic commands for an edit script, but adds some special copy commands. The script is computed separately for both the code and the data part of the object file, and merged afterwards. Some copy commands can be optimized that way. For dissemination, an algorithm called Ripple is used, that distributes the code packets to a selective number of nodes, not flooding the network. Corrupted or missing packets are retransmitted using a sliding window protocol, which allows the node to process or forward received packets while waiting for the retransmission of the missing packet.

2.3 Trickle

Trickle [4] is the epidemic algorithm used by Deluge for propagating and maintaining code updates in wireless sensor networks. A "polite gossip" policy is applied, where nodes periodically broadcast a code summary to the local neighbors, but stay quiet if they have recently heard a summary identical to theirs. A node that hears an older summary than its own broadcasts an update. Instead of flooding the network with packets, the algorithm controls the send rate so each node hears a small trickle of packets, just enough to stay up to date. An implementation of Trickle is contained in TinyOS 2.x.

2.4 SensorWare

In SensorWare [5], the developers set very high requirements on the hardware. It does not fit into the memory of popular sensor nodes and targets richer platforms to be developed in the future. In contrast to Maté, also complex semantics can be expressed. The program services are grouped into theme related APIs with Tcl-based scripts as the glue. Scripts located at various nodes use these services and collaborate with each other to orchestrate the data flow to assemble custom networking and signal processing behavior. Application evolution is facilitated through editing scripts and injecting them into the network. Both SensorWare and Maté can update application by replacing high-level scripts. They cannot permit the lower level binary code to be modified.

3. System Architecture

3.1 Network Topology

There are three types of node in a WASN: Leader Node, Function Node, and Sensor Node. They cooperate with each other to deal with necessary data, in order to achieve the goal of distributed computation and power consumption.

Fig. 1 shows adaptive SmartMote packet transmission in the network. User uses instructions defined by the system to set nodes behavior. The instructions will be compiled to byte codes by computer. It will encapsulate the byte codes into network packet and transmitted the packet to Leader Node. After Leader Node receiving and parsing the packet by SmartMote, it will distribute the packet to the Function Nodes in the network or execute the code itself. As the instructions we propose, they describe some behaviors affecting target nodes.

The packet described above is included instructions. When nodes send/receive packets, they will enable nodes to operate new behaviors. For an instance, if a packet describes the instruction of computing, nodes will compute sensing data base on the instruction after SmartMote parsing the packet. The data after dealing by the node will be passed to PC. Therefore, the distributed architecture enables data to distributed computation or update node's behavior. There are two issues supervening: 1) how to manage group nodes, and 2) how about the architecture of wireless sensor network.

a) Node Classification and Description

According to node hardware, capability, and electricity, two kinds of node are specified: Super Node and Sensor Node. Super Node provides data computing, coordination, and communicating. However, Sensor Node just collects necessary data and transmits it to Super Node or react the behavior that Super Node assigns it. Super Node is different from Sensor Node not only hardware specification but also inner component structure. Super Node enables real-time updating all code storage and altering it behavior. Otherwise, Sensor Node has a few algorithms hard-coded into each node but tunable through the transmission of parameters. On the other hand, Super Node can carry on Leader Node election. Leader Node is a cluster's head and the others are Function Nodes.

b) Leader Node Election and Inheritance

Each Super Node begins the status of the competition. After competition, one Leader Node and several Function Nodes will be identified. And then there is a table that records result of election in each node. If present Leader Node is destroyed or come to power of threshold limit value, backup scheme will be started to inherit to the leader. Hence, we are able to generate a new Leader Node efficiently and increase performance effectively.

The condition of inherit to leader is depended on the threshold limit value of power. When power of the present Leader Node come to threshold limit value, present Leader Node starts handing over its job to the next Function Node which ranks the first in the inheritance table. Finally, the new Leader Node broadcast update message to all nodes.

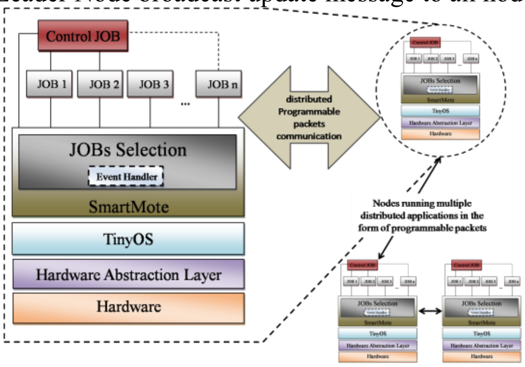


Fig. 2: Sensor framework

3.2 Sensor Framework

Fig. 2 shows SmartMote place inside a layered sensor node's framework. The lower layers are the raw hardware and the hardware abstraction layer (i.e., the device drivers). TinyOS exists on top of the low layers, which provides all the basic services and components of limited available resources that are needed by the layers above it. The SmartMote layer uses those functions and services offered by TinyOS to provide the run-time environment for the control JOBs. The layer for instance, includes event handler for events to register. The control JOBs rely completely on the SmartMote layer while populating through the network. Control JOBs use the native services that SmartMote provides as well as services provided by other JOBs to construct applications. Two things comprise SmartMote: 1) the language and 2) the supporting run-time environment.

a) The language and programming model

First, a language needs commands to act as the basic building blocks of the JOBs. These commands are essentially the interface to the abstraction services offered by SmartMote. Simple examples include: timer services, acquisition of sensing data, location discovery protocol. Second, a language needs constructs in order to tie these building blocks together in control JOBs. Some examples include: constructs for flow control, such as loops and conditional statements, constructs for variable handling and constructs for expressing evaluation. We call all these constructs the "net core" of the language, as they combine several of the basic building blocks to make actual control JOBs.

NesC [8], offering great modularity and portability, is considered as a suitable language for SmartMote. We choose the NesC core to be the net core in the SmartMote language. All the basic commands are defined as new NesC commands using the standard method that NesC provides for that purpose.

As SmartMote's design progressed over time, the set of commands changed considerably. We start with some basic commands and APIs for object mobility along with some commands for timer, network, and sensing abstraction, and kept adding commands as necessary. SmartMote declares, defines, and supports the creation of virtual devices. All abstraction services are represented as virtual devices. There is a fixed interface for all devices.

An intuitive description of a sensor node task (a part of a distributed application) has the form of a state machine that is influenced by external events. This is also the form of SmartMote JOBs. The programming model is as following: An event is described, and it is tied with the definition of an event handler. The event handler, according to the current state, will do some processing and possibly create some new events or/and alter the current state. For example, there is waiting for event a or b, c. if a device can produce events, a task is needed to accept event with waiting state that is waiting on the device's events. Although the JOBs are defining behavior at the node level, SmartMote is not a node-level programming language. It can be better viewed as an event-based language since the behaviors are not tied to specific nodes but rather to possible events that depend on the physical phenomena and the WASN state.

b) The run-time environment

Fig. 3 depicts abstracted view of SmartMote's run-time environment. Most of the threads running are coupled with a generic queue. Each thread "pends" on its corresponding queue, until it receives a message in the queue. When a message arrives it is promptly processed. The next message will be fetched, or if the queue is empty, the thread "pends" again on the queue. A queue associated with a JOB thread is receiving events (i.e., reception of network messages, sensing data, or expiration of timers). A queue associated with one of the three resource handling tasks, receives events of one type (from the specific device driver that is connected to), as well as messages that declare interest in this event type. For instance, the Sensing resource-handling task is receiving sensing data from the device driver and interests on sensing data from the JOBs. The JOB Manager queue receives messages from the network that wish to spawn a new JOB. There are also system messages that are exchanged between the system threads (like the ones that provide the Admission Control thread with resource metering information, or the ones that control the device drivers).

c) Code Transmitting and Updating

Fig. 4 shows the flow chart for a user using instruction to update sensor behavior. The instruction is translated to byte code and distributed to node in the network. When we want to update Function Node, Leader

Node will receive the byte code with updating instruction. Leader Node route the byte code to the Function Node according to its routing table. After the Function Node receiving the byte code, it parses the code by SmartMote and updates its behavior.

d) Programmable Packet Format

The described above is about how to generate a programmable packet and how to communicate and update in the WASN. We design a format of programmable packet. The programmable packet with executable program is shown as Fig. 5, in which the format of programmable packet. Restricted to TinyOS, only 29 bytes of packet length can be used. We set GID, STA, SRC, and DES to basic head. However, TYPE, LEN, and DATA/CODE describe about the information of packet.

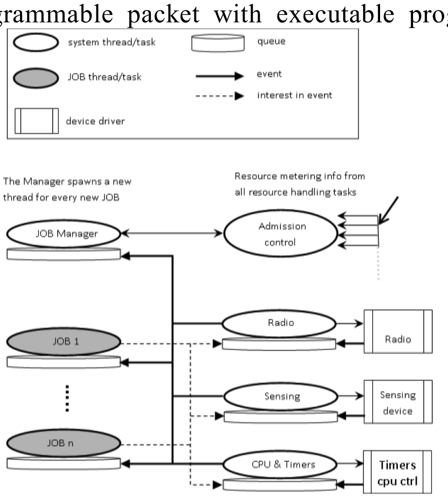


Fig. 3: Abstracted view of SmartMote's run-time environment

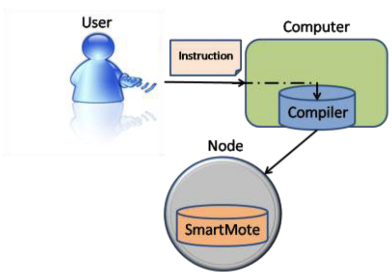


Fig. 4: User uses instruction to update sensor behavior.

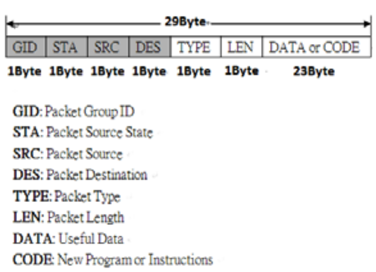


Fig. 5: Programmable packet format

4. SmartMote

SmartMote is a compact interpreter-like virtual machine [6] designed specifically for WASNs built on TinyOS. Instead of installing applications as binary objects on the sensor node, every node executes a byte code interpreter. SmartMote reads the special byte code commands from memory, and transforms these operations to TinyOS. Therefore reinstallation and rebooting are not required if the program is just some input data for this system. The flash memory size is 1024 KB. TinyOS, Code Store, and Data Store are allocated 128 KB. SmartMote, Register, and Temp Store are allocated 64 KB.

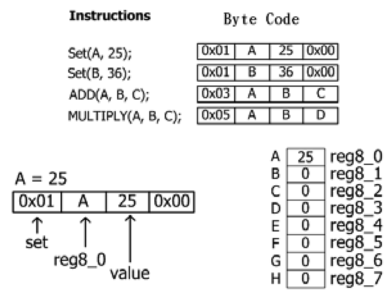


Fig. 6: An example of SmartMote instruction execution

4.1 SmartMote Architecture

In order to achieve the goal of real-time updating or computing, the loader is triggered to access the new code or parameters from flash memory. Then the Loader Node loads the byte code into SmartMote. Finally, SmartMote affects their behavior after parsing and executing the byte code.

4.2 SmartMote Instructions

The instructions are used by user to affect sensor behavior. There are four types of instructions: computing instruction, control instruction, system instruction, and network instruction. Fig. 6 shows an example of a SmartMote instruction execution. We compile instruction into byte code and write the 4-bits byte code to register. SmartMote parses and executes the byte code in order to affect behavior of a node.

5. Performance Analysis

In this chapter we present experiments as well as simulations on the performance of SmartMote. The experiments and measurements are conducted on a hardware platform, Octopus II [7]. On the other hand, for the simulations, we choose TOSSIM as the simulation platform.

5.1 Experimental Results and Analysis

The test bed is set to a 30 x 30m free space on our campus. We set four cases: 4, 8, 12, and 16 nodes in the free space. After sensor nodes update their behaviors, they send their sensing data to the leader node. Fig 7 shows the traditional cases with flooding method. When there are 16 nodes within 60 seconds of operation of the network, the calculated value is 3840. Form the results, we find the value of measurement is 3222 and its loss rate is 16%.

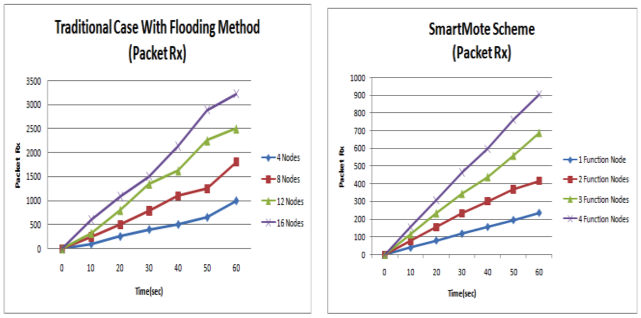


Fig 7&8: Traditional case with flooding method & Smart scheme (packet rx).

Fig 8 shows the result that uses SmartMote scheme. For each case that is considered in Fig 7, one function is generated for every four sensor nodes. For example, for the case of 16 nodes in Fig 7, four functions nodes are assigned. Therefore, for the traditional cases of 4, 8, 12, and 16 nodes (in Fig 7), we have 1, 2, 3, and 4 function nodes respectively (in Fig 8). In the SmartMote scheme, a leader node receives the result of computation from function nodes. However, each of function nodes manages 4 sensor nodes in the network. By the scheme, we are able to reduce packet collision and decrease the number of transmitted packets in the network. When receiving packets, SmartMote scheme is stable than flooding scheme.

TABLE 1: Power consumption parameters.

B_D	Number of data bytes in a packet
B_P	Number of preamble bytes in a packet
P_{Tx}	Power of transmitter in mA
P_{Rx}	Power of receiver in mA
P_r	Power of reading in flash ROM in mA
P_w	Power of writing in flash ROM in mA
$P_{Instruction}$	Power of instruction in mA
ζ	Constant determined by the MAC scheme and represents the average proportion of time spent in receive mode divided by that spent in transmit mode
ε	Power of waking up sensor node

a) Evaluation of Power Consumption

In our system, row data can be computed by SmartMote in sensor nodes. Processed data can be collected by function nodes and then transmitted to leader node. By the procedure, it decreases the amount of transmission and it also reduces the power consumption.

The following metric describes the power consumption of the transmitter and receiver node when updating Jobs and used this data to evaluate the physical layers of a sensor network. We model the power consumption P_{Total} of the reprogramming process with

$$P_{Total} = P_{Radio} + P_{FlashAccess} + P_{Computing} + P_{SensorStartup}$$

Where P_{Radio} is the power spent in transferring and receiving the JOB over the network, $P_{FlashAccess}$ the power cost of reading and writing the JOB in flash ROM, $P_{Computing}$ the power consumed by using instructions to compute data, and $P_{SensorStartup}$ the required power for waking up sensor node. P_{Radio}, $P_{FlashAccess}$, $P_{Computing}$, and $P_{SensorStartup}$ can be further extend to ($(B_D + B_P)(P_{Tx} + \zeta P_{Rx})$), $B_D(P_r + P_w)$, $B_D P_{Instruction}$, and ε.

TABLE 1 presents parameters of power consumption. Based on the structure and power consumption of each component, the value of P_{Total} can be written as

$$P_{Total} = (B_D + B_P)(P_{Tx} + \zeta P_{Rx}) + B_D(P_r + P_w) + B_D P_{Instruction} + \varepsilon$$

when $B_{D+} B_P \cong B_D$,

$$P_{Total} = B_D(P_{Tx} + \zeta P_{Rx} + P_r + P_w + P_{Instruction}) + \varepsilon$$

If there are k packets at most in the experiment, we can set the value of P_{Total} as

$$P_{Total} = \sum_{i=1}^{k} B_{D(i)}(P_{Tx} + \zeta P_{Rx} + P_r + P_w + P_{Instruction}) + \varepsilon$$

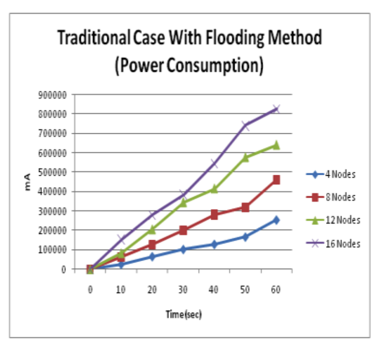


Fig 9: Traditional case with flooding method (power consumption).

In the formula above, we adopt the results from Fig 7 and 8 to evaluate power consumption in the WASN. For the example, we assume a 30x30m free space, ζ =1 [9], and ε =8.103mA. Fig 9 and 10 show the results of power consumption. In the same environment condition, we find that SmartMote scheme can reduce up to 72% power consumption.

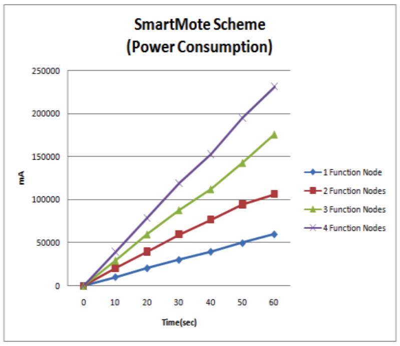


Fig 10: SmartMote scheme (power consumption).

5.2 Simulations

a) Results

Assuming the leader node transmits a job to 100 sensor nodes for updating their behavior. Sensor nodes base on the job to react to the behavior and send sensing data to leader node through function nodes.

Fig 11 shows the distribution of completion time for individual nodes. All nodes have a completion time bigger than 1 second, but less than 15 seconds. We note that the average range of completion time for an individual is 0.15± 0.05 second. Fig 12 presents the final results revealing that the SmartMote scheme is considerably faster than the traditional cases with flooding scheme. Furthermore, the bulk of the delay in two schemes shows that flooding scheme spent more time in the communication part and computation part than SmartMote scheme. In flooding scheme, because its packet lost rate is higher than SmartMote scheme, more retransmission is required. Moreover, flooding scheme needs centralized computation in the leader node, so it also spends much time in computation part.

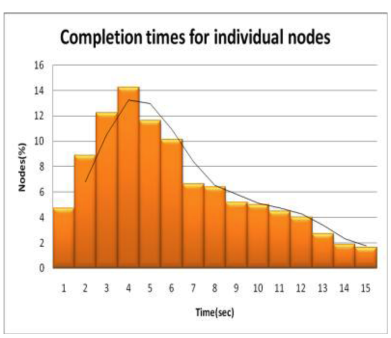


Fig 11: The distribution of completion times for individual nodes.

6. Conclusions

In this paper, an application to update WASNs with programmable packet and SmartMote is designed and implemented for TinyOS on the SmartMote platform. We present our framework for dynamic and efficient WASN programmability. Through our implementation we are able to achieve the goals of power consumption and behavior update. SmartMote system makes WASN platforms open to transient users with dynamic needs. This fact, apart from giving an important flexibility advantage to deployed systems, greatly facilitates researchers to cross the simulation boundary and start testing their algorithms/protocols in real platforms.

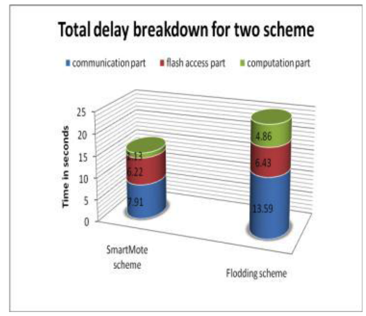


Fig 12: Total delay breakdown for two schemes.

Acknowledgement

This study is conducted under the "Applied Information Services Development & Integration project" of the Institute for Information Industry which is subsidized by the Ministry of Economy Affairs of the Republic of China.

6.1 References

[1] J. Hill, R. Szewczyk, A. Woo, S. Hollar, D. Culler, K. Pister, "System Architecture Directions for Networked Sensors," ASPLOS-IX proceedings, Cambridge, MA, USA, Nov. 2000

[2] Crossbow Technology, "Mote in Network Programming User Reference," TinyOS document, http://webs.cs.berkeley.edu/tos/tinyos-1.x/doc/Xnp.pdf.

[3] T. Stathopoulos, J. Heidemann, and D. Estrin, "A Remote Code Update Mechanism for Wireless Sensor Networks," CENS Technical Report No. 30, Nov. 2003

[4] P. Levis, N. Patel, D. Culler, and S. Shenker, "Trickle: A Self-Regulating Algorithm for Code Propagation and Maintenance in Wireless Sensor Networks," NSDI'04, pp. 15-28, 2004

[5] A. Boulis, C.-C. Han, and M.B. Srivastava, "Design and Implementation of a Framework for Efficient and Programmable Sensor Networks," MobiSys'03 Proceedings, pp. 187-200, New York, NY, USA, 2003

[6] P. Levis and D. Culler, "Mate: A Tiny Virtual Machine for Sensor Networks," ASPLOS X Proceedings, 2002

[7] Moteiv, "Tmote Sky Data Sheet," 2006

[8] D. Gay, P. Levis, R. von Behren, M. Welsh, E. Brewer, D. uller, "The nesC Language: A Holistic Approach to Networked Embedded Systems," ACM PLDI'03, San Diego, CA, USA, Jun. 2003

[9] J. A. a. J. Rabaey, "The Energy-per-Useful-Bit Metric for Evaluating and Optimizing Sensor Network Physical Layers."

A RSSI-based Algorithm for Indoor Localization Using ZigBee in Wireless Sensor Network

Yu-Tso Chen[1], Chi-Lu Yang[1,2], Yeim-Kuan Chang[1], Chih-Ping Chu[1]
[1]Department of Computer Science and Information Engineering, National Cheng Kung University
[2]Innovative DigiTech-Enabled Applications & Service Institute, Institute for Information Industry
[1]Tainan, Taiwan R.O.C. [2]Kaohsiung, Taiwan R.O.C.
{ p7696147, p7896114, ykchang, chucp}@mail.ncku.edu.tw

Abstract

For the various applications in home automation, the service system requires to precisely estimate user's locations by certain sensors. It is considered as a challenge to automatically serve a mobile user in the house. However, indoor localization cannot be carried out effectively by the well-know Global Positioning System (GPS). In recent years, Wireless Sensor Networks (WSNs) are thus popularly used to locate a mobile object in an indoor environment. Some physical features are widely discussed to solve indoor localization in WSN. In this paper, we inquired about the RSSI solutions on indoor localization, and proposed a Closer Tracking Algorithm (CTA) to locate a mobile user in the house. The proposed CTA was implemented by using ZigBee CC2431 modules. The experimental results show that the proposed CTA can accurately determine the position with error distance less than 1 meter. At the same time, the proposed CTA has at least 85% precision when the distance is less than one meter.

Keywords: indoor localization, home automation, ZigBee modules, wireless sensor networks

1. Introduction

For a large number of applications in home automation, the service system requires to precisely sensing user's locations by certain sensors. Moreover, the system sometimes requires recognizing the time and the weather for making decisions. On the other hand, the users always hope to be served correctly and suitably by the service system in the house. For satisfying the users' demands, one of the most key successful factors is to accurately estimate the user's location. It is considered as a challenge to automatically serve a mobile user in the house.

Indoor localization cannot be carried out effectively by the well-know Global Positioning System (GPS), which is subject to be blockaded in the urban and indoor environments [1-4]. Thus in recent years, Wireless Sensor Networks (WSNs) are popularly used to locate mobile object in the indoor environment. Some physical features are widely discussed to solve indoor localization in WSN. Received signal strength indication (RSSI) is the power strength of radio frequency in a wireless environment. The RSSI value can be regularly measured and monitored to calculate distance between objects. Time of arrival (TOA) means the travel time of a radio signal from one single sender to another remote receiver. By computing the signal transmission time between a sender and a receiver, the distance could approximately be estimated. Time difference of arrival (TDOA) is computed based on the emitted signals from three or more synchronized senders. It also refers to a solution of locating a mobile object by measuring the TDOA.

In this paper, we inquired about the RSSI solutions on indoor localization, and proposed a new RSSI-based algorithm and implemented it using ZigBee CC2431 modules in wireless sensor network. The rest of this paper is organized as follows. In Section 2, we briefly introduce the related work on indoor localization in WSN. In Section 3, we first define relevant arguments to describe our algorithm. We then carefully explain the proposed algorithm. In Section 4, the experimental results are analyzed and discussed to validate our algorithm. We show our algorithm is more accurate by comparing with the others methods. The conclusion and future work of the study are summarized in Section 5.

2. Related Work

ZigBee solutions are widely applied in many areas, such as home automation, healthcare and smart energy *(ZigBee Alliance)*. ZigBee is a low-cost, low-power, low data rate and wireless mesh networking standard originally based on the IEEE 802.15.4-2003 standard for wireless personal area networks (WPANs). The original IEEE 802.15.4-2003 standard has been superseded by the publication of IEEE 802.15.4-2006 for extending its features [5, 14]. While many techniques related to ZigBee have also been applied to indoor localization, we choose to focus on 2-dimension localization issues for the following introduction.

2.1 Fingerprinting

The Fingerprinting (FPT) systems are built by analyzing the RSSI features. The RSSI features are pre-stored in a database and are approximately retrieved to locate a user's position [8-11]. The key step of FPT is that the blind node is put at pre-defined anchor positions in advance. By RSSI, the blind node continuously sends requests to its surrounding reference nodes and receives responses from these reference nodes. The FPT system can then continuously record these responses to analyze its features until the analyzed results are characteristically stable. In general, different anchors should be distinct from different RSSI features. In FPT, the mobile object is approximately located by comparing the current RSSI with the pre-stored RSSI features.

Denote a series offline training measurement of reference node k at location L_{ij} is $L=[l_{ij}^{k0},...,l_{ij}^{kM-1}]$ which enables to compute the histogram h of RSSI.

$$h_{ij}^{k}(\zeta)=\frac{1}{M}\sum_{m=0}^{M-1}\delta(l_{ij}^{km}-\zeta),-255\le\zeta\le 0 \qquad (1)$$

The reference nodes are indexed with k. The parameter δ represents the Kronecker delta function [8, 11].

2.2 Real-Time Tracking

The method, which can locate a mobile object by at least three reference nodes without pre-trained database, is named Real-Time Tracking (RTT) [1-4, 6-7]. The RTT System can convert the RSSI to a distance by specific formulas. Trilateration is a method to determine the position of an object based on simultaneous range measurements from at least three reference nodes at known location [1]. Trilateration requires the coordinates of at least three reference nodes (X_i, Y_i) and the distances d_p^i between the blind node and the pre-positioned reference nodes. The target's position $P(X_p, Y_p)$ can be obtained by MMSE [3]. The difference between actual and estimated distance is defined by formula *(2)* where i is a reference position and p is a mobile object.

$$d_P^i=\sqrt{(x_i-x_p)^2+(y_i-y_p)^2} \qquad (2)$$

Eq. (2) can be transformed into

$$(d_P^i)^2=(x_i-x_p)^2+(y_i-y_p)^2 \qquad (3)$$

Then Eq. (3) is able to be transformed into

$$\begin{bmatrix}(d_P^1)^2-(d_P^2)^2+(x_2^2+y_2^2-x_1^2-y_1^2)\\(d_P^1)^2-(d_P^3)^2+(x_3^2+y_3^2-x_1^2-y_1^2)\\\dots\\(d_P^1)^2-(d_P^N)^2+(x_N^2+y_N^2-x_1^2-y_1^2)\end{bmatrix}=\begin{bmatrix}2(x_2-x_1) & 2(y_2-y_1)\\2(x_3-x_1) & 2(y_3-y_1)\\\dots & \dots\\2(x_N-x_1) & 2(y_N-y_1)\end{bmatrix}\begin{bmatrix}x_p\\y_p\end{bmatrix} \qquad (4)$$

Therefore, Eq. (4) is transformed into Eq. (5), which can be solved using the matrix solution given by Eq. (6). Position $P(X_p, Y_p)$ can be obtained by calculating Eq. (6).

$$b=A\begin{bmatrix}x_p\\y_p\end{bmatrix} \qquad (5) \qquad\qquad \begin{bmatrix}x_p\\y_p\end{bmatrix}=(A^TA)^{-1}*(A^Tb) \qquad (6)$$

Where

$$b=\begin{bmatrix}(d_P^1)^2-(d_P^2)^2+(x_2^2+y_2^2-x_1^2-y_1^2)\\(d_P^1)^2-(d_P^3)^2+(x_3^2+y_3^2-x_1^2-y_1^2)\\\dots\\(d_P^1)^2-(d_P^N)^2+(x_N^2+y_N^2-x_1^2-y_1^2)\end{bmatrix} \qquad (7)$$

$$A=\begin{bmatrix}2(x_2-x_1) & 2(y_2-y_1)\\2(x_3-x_1) & 2(y_3-y_1)\\\dots & \dots\\2(x_N-x_1) & 2(y_N-y_1)\end{bmatrix} \qquad (8)$$

3. Proposed Algorithm

3.1 Definitions

A blind node refers to a mobile object. A reference node is a fixed node that responds its RSSI to assist locating the blind node. In this study, both the blind node and the reference node are ZigBee modules. In order to describe our proposed algorithm, the following terms are principally defined. These terms are categorized into primitive terms, original physical terms and derived terms. The primitive terms are defined as follows:

$N_{neighbor}$ = the number of reference nodes which close to blind node within one hop currently

BID = a pre-defined identification of a blind node, which is a mobile object.

RID = a pre-defined identification of a reference node (a fixed object), where $1 \le RID \le N_{neighbor}$

$R_{threshold}$ [RID][d] = the RSSI of RID within the pre-defined threshold at distance d, where distance d is a set = {d(m) | 0.5, 1, 1.5, 2.0, 2.5, 3.0}

M_{ACA} = the mode of approximately closer approach for Tracking (the improved algorithm)

M_{RTT} = the mode of Real-Time Tracking

The values of RSSI thresholds of *RID* within distance *d* are pre-trained and stored in the database. The terms of physical arguments, which are originally received from ZigBee blind node, are defined as follows:

$R_{now}(x)$ = the current value of the measured RSSI of x, where variable x refers to RID

rid = an index of R_{now}, where rid < $N_{neighbor}$

The derived terms, which values are calculated from the physical terms and primitive terms, are defined as follows:

CloserList[x] = a list RID of sorted by $R_{now}(x)$, where $R_{now}(x)$ within $R_{threshold}[x][d]$ and $R_{now}(x) \le R_{now}(x-1)$, $1 \le x \le N_{neighbor}$

SortedList[x] = a list RID of sorted by $R_{now}(x)$, where R_{now} (SortedList[x]) $\le$ *R_{now}* (SortedList[x-1])

ClosestRID = a rid refers to RID, which is the closest node near the blind node (the mobile object; BID), and where R_{now} (ClosestRID) is within $R_{threshold}$

C_R = a record for tracking the mobile object

M_C = Current localization mode = {M_C | M_{ACA}, M_{RTT}}

3.2 Closer Tracking Algorithm

The locating style of the FPT has its own specific advantage and disadvantages, while the RTT style also has its own. The features of the two styles are characteristically complementary. Therefore, we proposed a compound algorithm to determine the usable mode at suitable time. Furthermore, we improved the FPT algorithm at the same time. This idea is also emerged from our observation on elder persons in the house. The elders usually stay on the same positions, such as sofa, table, water cooler or bed. They even frequently stay in front of the television or near the door for a long time. The time they are moving is much less than they are staying, while they are in their house. Since we look forward to provide automatic applications suitably for elders in their house, we can ideally design a position tracking algorithm based on above observation. The proposed algorithm for closer tracking (CTA) was specifically designed to improve the automatic applications. The CTA is carried out by the following four steps.

Step1 – [Build Neighbor List]

The blind node *BID* (the mobile object) periodically receives RSSI (R_{now}) from its neighbor nodes (*RIDs*) by broadcasting its requests. The neighbor nodes will be recorded by comparing their RSSIs with the pre-defined thresholds ($R_{threshold}$). In other words, if the RSSI of the *RID* is within the $R_{threshold}$ at *distance d*, the *RID* will be stored into the *CloserList.*

Step2 – [Determine Mode]

If there are records stored in *CloserList*, the improved FPT will be executed to locate the mobile object. In other words, if there is no record in the *CloserList*, the RTT will be executed for locating the mobile object.

Step3 – [Adapt Assistant Position]

It's likely that there is only one record in the *CloserList*. If the special situation occurs, we should need an extra data structure - *SortedList*. The *SortedList* is an array used to store the ordering *RIDs*, which are sorted by the received RSSIs. Nevertheless, the closest *RID* (*ClosestRID*) should not be stored into the *SortedList*. In next step, the *CloserList* and *SortedList* will be used to locate the mobile object more precisely under M_{ACA} mode.

Step4 – [Approximately Closer Approach]

The improved FPT, which is named approximately closer approach (ACA), is divided into two phases. In the first phase, *ClosestRID* is used to figure out a circular range, since the RSSI of the *ClosestRID* is within the pre-defined threshold at distance *d*. The plane of *ClosestRID* range can be conceptually divided into four sub-planes. In the second phase, the *RIDs* in the *CloserList* will be iteratively retrieved to select the sub-planes for narrowing down the outer range. For example, let's assume the *CloserList = {Ref4, Ref1, Ref5}* and *ClosestRID = Ref3*. In Fig. 1, a virtual circle surrounding the node *Ref3* will be first figured out, since the *ClosestRID* refers to *Ref3*. The plane of *Ref3* range can be conceptually divided into four sub-planes, such as *R1*, *R2*, *R3* and *R4*. In the second round, the sub-plane *R2* will be selected since the *Ref4* is the first *RID* in the *CloserList*. The other *RIDs* in the *CloserList* will be iteratively selected to narrow down the range. The iteration will be stopped until the *CloserList* is empty. The pseudo codes of the CTA, which contains the ACA, are showed in Table 1.

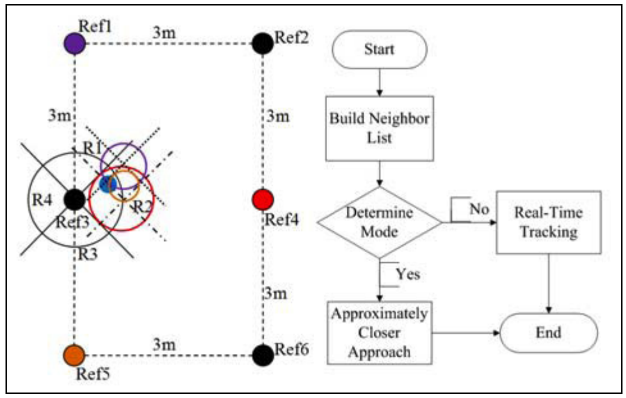


Fig. 1 Concept and flow of the Proposed Algorithm

Table 1 the pseudo codes of the CTA

```
Algorithm_Closer_Tracking(int *Rnow)
{
//////Initial/////////////////////////////////////////////////////////////////////////
short CloseList [8]={-1};
int k=0;
const int row=3;
const int col=2;
//////Step1 – Build Neighbor List////////////////////////////////////////////
for (dis = 0.5 ; dis <= 2.0 ; dis += 0.5){
    for (rid = 1 ; rid <= Nneighbor ; rid++){
        if (Rnow[rid] within Rthreshold[rid][dis]){
            CloseList[k] = rid;
            k++;
        }//end if
    }//end for
} //end for loop
//////Step2 – Determine Mode/////////////////////////////////////////////////
if (k == 0) { // No record in the CloserList
    MC = MRTT
    break; //Change to Real-Time Tracking Mode
} //end if
//////Step3 – Adapt Assistant Position/////////////////////////////////////
//////Only ClosestRID in the CloserList//////////////////////////////////
if (k == 1) {
    for (int x = 1; x < Nneighbor ; x++){
        CloseList[x] = SortedList[x-1];
    } //end for
    k = Nneighbor;
} // end if
//////Step4 –Approximately Closer Approach//////////////////////////////
ClosestRID = CloseList[0];
for (int s = 0 ; s < k ; s++){    //FPT
    switch (CloseList[s+1] - ClosestRID){
        case 1:
            CR[s] = R2;     break;
        case -1:
            CR[s] = R4;     break;
        case col:
            CR[s] = R3;     break;
        case -col:
            CR[s] = R1;     break;
        default:    //other 4 direction
    } //end switch
} //end for
MC = MACA
} //end Closer Tracking Algorithm
```

4. Implementation and Experiment

The ZigBee modules are used in this experiment. The CC2431 chip stands for the blind node and the CC2430 chips stand for the reference nodes. The specific features of these chips are listed in Table 2, and the figure of CC2431 is showed in Fig. 2. The RSSI values are long-term measured in the experiment, and all the values are stored in a database for further analysis. The proposed CTA is programmed by using the C#.NET language.

4.1 Findings

We measured 1-D RSSI in different environments, which electromagnetic waves are isolated, absorbed or normal. In Fig. 3, the x-axis represents the various distances between a blind node and a reference node, such as 0.5, 1, 1.5, 2.0, 2.5 and 3 meters. The y-axis represents the measured RSSI values. The RSSI values are measured until the statistic results are stable. In order to observe the data, all the measured values are added by one hundred. The statistic results and the standard deviation σ of the stable RSSI are shown in Fig. 3. The σ values are further utilized to define the thresholds.

The following formula provided by Texas Instruments (TI), which represented the relationship between RSSI and the estimated 1-D distance, is shown as follows:

$$RSSI = -(10n\log_{10} d + A) \qquad (9)$$

While n is a signal propagation constant or exponent, d is a distance from the blind node to the reference node and A is the received signal strength at 1 meter distance. According to the formula (9), the 1-D distance d can be derived from the measured RSSI values of Fig. 3 and shown in Fig. 4.

Table 2 Features of CC2431

Features	Values
Radio Frequency Band	2.4GHz
Chip Rate(kchip/s)	2000
Modulation	Q-QPSK
Bit rate(kb/s)	250
Sensitivity	-92dBm
Data Memory	8KB
Program Memory	128KB internal RAM
Spread Spectrum	DSSS

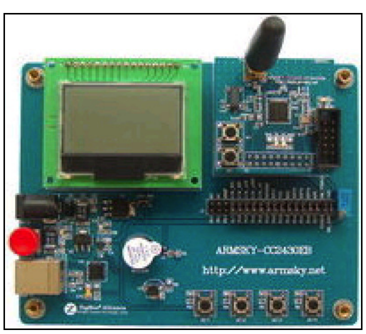


Fig. 2 CC2431 module

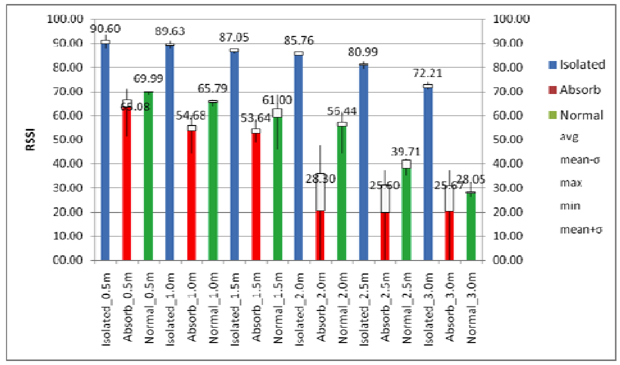


Fig. 3 RSSI thresholds

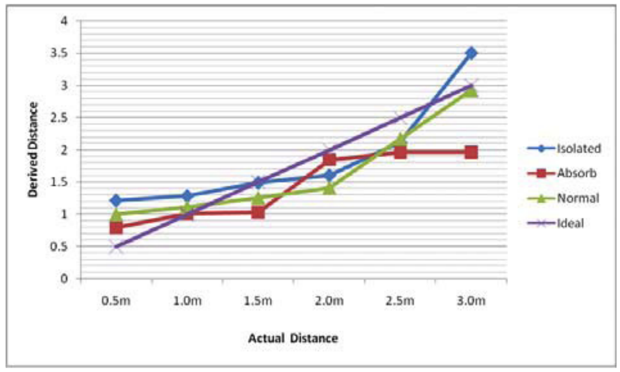


Fig. 4 Actual distance and derived distance (A, n) with Isolated (6, 4); Absorb (45, 10); Normal (30, 9)

4.2 Experimental Results

In this experiment, an actual position is represented by the coordinate (x, y), and an estimated position is represented by the coordinate (i, j). Therefore, we can simply define the accurate distance and represent by an Error Distance formula as follows:

$$Dist.(L_{xy}, L_{ij}) = \sqrt{(x-i)^2 + (y-j)^2} \qquad (10)$$

In order to validate accuracy of the proposed CTA, we implemented and compared the proposed CTA with the FPT [9] and RTT [12], which are experimented by using the CC2431 location engine. The experimental results are shown in Fig. 5 and Fig. 6. The x-axis represents the distance from the blind node to the closest reference node. The y-axis represents the difference between an actual position and the estimated position.

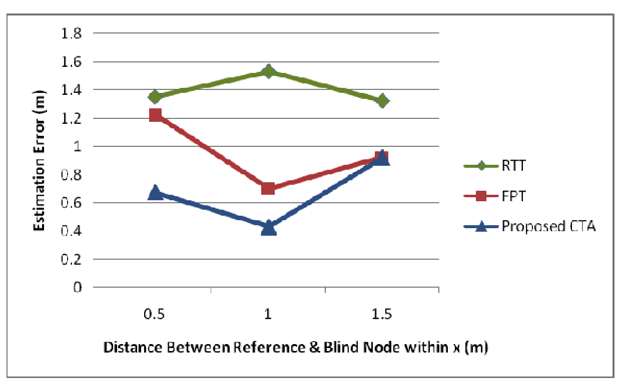


Fig. 5 Estimation errors at distance {0.5, 1.0, 1.5} meters (Accuracy)

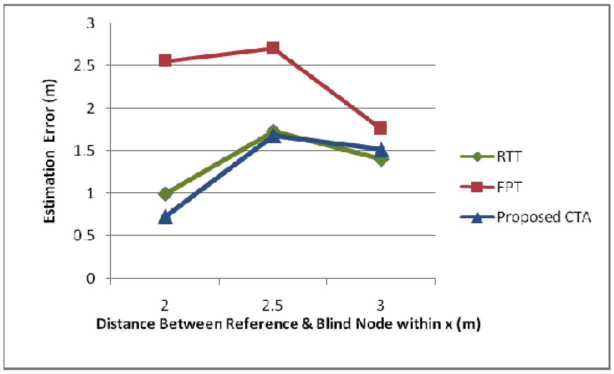


Fig. 6 Estimation errors at distance {2.0, 2.5, 3.0} meters (Accuracy)

As we can see from the experimental results in Fig. 5, when a blind node approaches to any reference node, our algorithm can accurately determine the position with error distance less than 1 meter. The accuracy of the CTA is better than the other methods. At the same time, the FPT method is accurate enough when the blind node is moving close to the pre-trained positions. Furthermore, the estimation errors calculated by CC2431 are quite stable in Fig. 5, and the accuracy of RTT method is quite independent of the positions of the reference nodes.

In Fig. 6, the distances from the blind node to the closest reference node are increased. Therefore, the RSSI values are more interfered by background noise, and the variances are increased. In FPT method, the signal features are diminished, so that the estimation errors are obviously increased. In other words, the FPT method cannot determine the position accurately when the distance from the blind node to the closest reference node is more than two meter. Under this condition, our proposed CTA changed the operational mode from the ACA to the RTT mode. As a result, the accuracy of the proposed method is close to those of the RTT method. In the case of x = 2.0m, the proposed CTA is slightly more accurate than the RTT method. In the other case of x = 3.0m, the proposed CTA is slightly worse than the RTT method.

In Fig. 7 and Fig. 8, we show the precision of the proposed CTA, the FPT, and the RTT. The precision is defined as follows:

$$\frac{Number\_of\_within\_Acceptable\_Error\_Distance}{Total \quad Estimated \quad Times} \qquad (11)$$

For the experimental design in Fig. 7, the acceptable error distance is set as 1 meter. Under this condition, the estimation errors, which values are less than or equal to 1 meter, are selected to calculate precision. As we can see, the proposed CTA has at least 85% precision when the distance is less than one meter. The CTA has higher precisions than the other methods. In Fig. 8, the precision is low yet in the case of x=2.5. This is because that most estimated errors stay in the range of 1.5 and 1.8. That's an interesting situation.

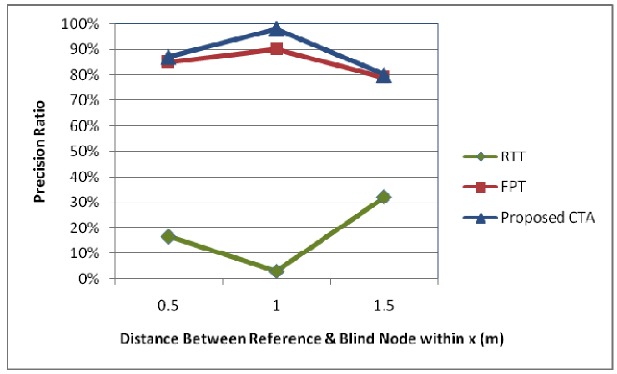


Fig. 7 Precision when error distance within 1.0 m

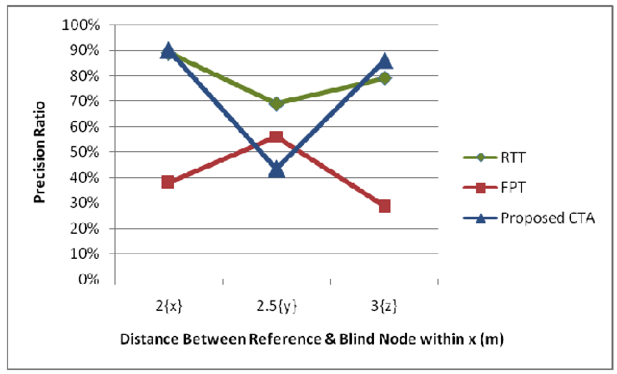


Fig. 8 Error distance {x} within 1.3m; Error distance {y, z} within 1.7m

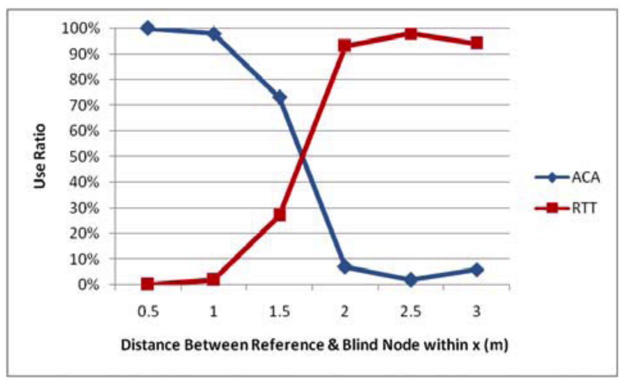


Fig. 9 Usage ratio of ACA & RTT modes

We showed the mode-changed functionality of the proposed CTA at various distances. The usage ratios of the ACA and the RTT are displayed in Fig. 9. As we can see, the ACA method is useful if the distance is less than 1.5 meters. Furthermore, the ACA mode will be changed to RTT if the distance increases over 1.5 meter. The mode-changed operation can be practically made according to the threshold we set. As a result, the proposed CTA can select an adaptive mode to obtain more precise location. The usage ratios of ACA and RTT are showed in Fig. 9.

5. Conclusion and Future Work

In this paper, we inquired about the RSSI solutions on indoor localization, and proposed a new RSSI-based algorithm using ZigBee CC2431 modules in wireless sensor network. Moreover, we improved the FPT algorithm at the same time. The mode-changed operation of the proposed CTA is even designed for combining the improved FPT and the RTT methods. The functionality can adapt the operational modes according to the thresholds, which we set and mentioned in the findings. As a result, the proposed CTA can suitably select an adaptive mode to obtain more precise locations. The experimental results show that the proposed CTA can accurately determine the position with error distance less than 1 meter. At the same time, the proposed CTA has at least 85% precision when the distance is less than one meter.

For the various applications in home automation, the proposed CTA can be applied to provide correct and suitable services by estimating user's locations precisely. In the future, the proposed CTA can even bring promising quality of services on caring elders in the house. At the same time, we will try to improve the real-time tracking algorithm of the CTA for increasing the accuracy of the uncovered ranges, which positions are beyond the reference nodes.

Acknowledgements

This research was partially supported by the second Applied Information Services Development and Integration project of the Institute for Information Industry (III) and sponsored by MOEA, Taiwan R.O.C.

References

[1] Erin-Ee-Lin Lau, Boon-Giin Lee, Seung-Chul Lee, and Wan-Young Chung, "Enhanced RSSI-Based High Accuracy Real-Time User Location Tracking System for Indoor and Outdoor Environments," International Journal on Smart Sensing and Intelligent Systems, Vol. 1, No. 2, June 2008.

[2] Youngjune Gwon, Ravi Jain, and Toshiro Kawahara, "Robust Indoor Location Estimation of Stationary and Mobile Users," in Proc. of IEEE INFOCOM, March 2004.

[3] Masashi Sugano, Tomonori Kawazoe, Yoshikazu Ohta, and Masayuki Murata, "Indoor Localization System Using RSSI Measurement of Wireless Sensor Network Based on ZigBee Standard," in Proc. of Wireless Sensor Networks 2006 (WSN 2006), July 2006.

[4] Stefano Tennina, Marco Di Renzo, Fabio Graziosi and Fortunato Santucci, "Locating ZigBee Nodes Using the TI's CC2431 Location Engine: A Testbed Platform and New Solutions for Positioning Estimation of WSNs in Dynamic Indoor Environments," in Proc. of the First ACM International Workshop on Mobile Entity Localization and Tracking in GPS-less Environments (MELT 2008), Sep. 2008.

[5] IEEE 802.15 WPAN™ Task Group 4, IEEE 802.15.4-2006, http://www.ieee802.org/15/pub/TG4.html

[6] Allen Ka, and Lun Miu, "Design and Implementation of an Indoor Mobile Navigation System," Master thesis of CS at MIT 2002.

[7] P. Bahl and V. Padmanabhan, "RADAR: An In-Building RF-based User Location and Tracking System," In Proceedings of the IEEE INFOCOM 2000, March 2000, pp. 775-784.

[8] Angela Song-Ie Noh, Woong Jae Lee, and Jin Young Ye, "Comparison of the mechanisms of the ZigBee's indoor localization algorithm," in Proc. SNPD, pp.13-18, 2008.

[9] Qingming Yao, Fei-Yue Wang, Hui Gao, Kunfeng Wang, and Hongxia Zhao, "Location Estimation in ZigBee Network Based on Fingerprinting," in Proc. IEEE International Conference on Vehicular Electronics and Safety, Dec 2007.

[10] Shashank Tadakamadla, "Indoor local positioning system for zigbee based on RSSI", M.Sc. Thesis report, Mid Sweden University, 2006.

[11] C. Gentile and L. Klein-Berndt, "Robust Location Using System Dynamics and Motion Constraints", in Proc. of the 2004 IEEE International Conference on Communications, vol. 3, pp. 1360-1364, June 2004.

[12] System–on–chip for 2.4 GHz ZigBee /IEEE 802.15.4 with location engine., Datasheet, Texas Instruments, http://focus.ti.com/lit/ds/symlink/cc2431.pdf, July 2007.

[13] K. Aamodt., CC2431 Location Engine. Application Note AN042, Texas Instruments.

[14] ZigBee Alliance, ZigBee Specification Version r13, San Ramon, CA, USA, Dec. 2006.

A Personalized Service Recommendation System In a Home-care Environment

Chi-Lu Yang[1,2], Yeim-Kuan Chang[1], Ching-Pao Chang[3], Chih-Ping Chu[1]
[1] Department of Computer Science and Information Engineering, National Cheng Kung University
[2] Innovative DigiTech-Enabled Applications & Service Institute, Institute for Information Industry
[3]Department of Information Engineering, Kun Shan University
[1,3]Tainan, Taiwan R.O.C.
[2]Kaohsiung, Taiwan R.O.C.
stwin@iii.org.tw, {ykchang, chucp}@mail.ncku.edu.tw, chingpao@ms28.hinet.net

Abstract

Many bio-signals of the chronic patients could be measured by various bio-devices and transferred to back-end system over the wireless network through the homebox. In a home-care environment, it becomes more complex to reliably process transmitting and receiving these bio-signals by the homebox. While the bio-devices increasing, the process is even much more complex. In addition, the chronic patients always hope to be served correctly and suitably by the service system in the house. Therefore, we have to provide services such as adjusting room temperature and lighting etc to make those patients' daily lives easy. In this paper, we propose a personalized service recommendation system (PSRS) based on users' preferences and habits. The PSRS has capability of providing suitable services. Furthermore, we construct personal models to record the patients' daily activities and habits. Through the models, the system will be able to automatically launch to safety alert, recommendable services and healthcare services in the house. In the future, the proposed system and models could even carry out the mobile health monitor and promotion in a home-care environment.

Keywords: personalized service, home-care system, service-oriented architecture

1. Introduction

Many bio-signals of the chronic patients could be measured and transferred over the wireless network through the homebox [1], [18]. However, bio-devices are gradually increasing recently. To manage the devices is thus becoming much more complex. Due to this situation, performance level is seen to decrease when a large number of data change occurs. In addition, the chronic patients always hope to be served correctly and suitably by the service system in their houses. Therefore, we have to provide services such as adjusting room temperature and lighting etc to make those patients' daily lives easy. If the system could actively predict the patient's preferences or habits, the system will be able to serve the person in advance with high quality of service.

In this paper, we developed a personalized service recommendation system based on patient's preferences in a home-care environment. For the recommendation services, we first explored the process of generating the recommendable service. We then constructed personal models by analyzing the patient's activity patterns. Through the personal models, the system will be able to automatically launch to safety alert, recommendable services and healthcare services in the house. The proposed system and models could even carry out the mobile health monitor and promotion. The rest of this paper is organized as follows. In Section 2, we are going to introduce the recommendation and personalization services. In Section 3, the proposed system and service groups are described. The processes of generating recommendable services are also mentioned. In Section 4, the personal models are carefully explained. The experiments and test cases are discussed in Section 5. The conclusion of the study is summarized in Section 6.

2. Related Work

2.1 Recommendation Service

The recommendable services are popularly applied on the Internet, such as the on-line recommended services (amazon.com), customized services (mywashington post.com), personalized advertisements (yesmail.com, yahoo.com) and other similar services [2]. By retrieving and analyzing the interactions between the users and the systems, recommendations services could be precisely delivered. The recommendable services are sometimes generalized to match the personal preferences [3]. In order to fit in with the user's demands, services are personalized and recommended based on the user's preferences and the contexts. Studies in applied systems showed that recommendations based on the user's habits can friendly get user's responses [4]-[6]. These results fit in with the studies in human-computer interaction and e-learning domains.

The Recommendation system could even provide custom-oriented services which differ from traditional service system. The services of the system could be personalized according to personal profiles. In order to achieve this goal, the primary step is to collect various information sources. These sources could be approximately classified into two types. The first one is user-relevant information, such as name birthday, health status, habits, and behavior patterns. The second type comes from the environment, such as the statuses of the devices, interactions between users and devices, the weather, the time, temperature, brightness and the others. The two kinds of these sources are primary foundations to build personalization services. Unfortunately, some sources are dynamically changed by the external factors. Furthermore, the user's demands are even too diverse to be monitored effectively. As a result, it is a challenge to recommend suitable services to a user.

2.2 Personalization Service

If we wish to properly recommend the services to a user, we should not only pay attention on the data sources. We also have to concentrate on personalization. For the service personalization, the key factors are to sensing the user's preferences and habits. Through these personal patterns, the existed services could be possibly adapted to match the user's needs. A recommendable service system could be inquired by the following viewpoints: user modeling, context modeling, semantic interoperability and service composition, self-service management, and so on. Using user models to predict user's needs is one of the most popular methods. An excellent user model would be able to select the proper attributes for exploring the user's behavior patterns [7]. The recommendable services could be dynamically composed and properly provided based on the users' patterns in specific environments.

The overlay model is a modeling technique based on collecting user's behaviors [8]. The primary idea of the overlay model is that a user's behavior is a subset of all users' behaviors. Therefore, a common model could be built by generalizing all users' behaviors. Individual model could then be established by comparing it with the common model.

The stereotype users' model is a speedy modeling technique. The model could be fundamentally built up, even when it lacks the user's behaviors component [9] -[10]. Although the model is built by the approximate value, it could perform effectively in many applications [11]. In order to build the stereotype model, the following elements are needed: user subgroup identification (USI), identification of users' key features (IUF), and representation template (RT). The first element is used to identify the subgroups' features. Users in a subgroup have application-relevant features on their behavior models. The second element is used to define the users' key features, which differ from the other subgroups. Furthermore, the presence and absence features should be clearly identified for decision support. The third element is hierarchically represented. The representations should be distributed in different systems. The representation templates in subgroups are named as stereotypes. The hierarchical style could precisely describe the user's behaviors when it goes down to low hierarchical nodes.

2.3 Interface Management and Query

Interface management is a mechanism for managing and providing various services to others. Web service is one of the most popular techniques nowadays. Through web services, various services are distributed in different systems and managed by individuals. The World Wide Web Consortium proposed three major roles for web services. (1) Service provider is defined to provide remote services. (2) Service registry is defined to provide registration and publication. (3) Service consumer is a role which requests to serve and receive the services. First, the service provider generates service descriptions and registers them into the service registry. The service consumer then requests the service registry and receives the interface descriptions. The relevant techniques are Web Services Description Language (WSDL) [13], Simple Object Access Protocol (SOAP; [14]), Universal Description, Discovery, and Integration (UDDI) [15], and Extensible Markup Language (XML) [12].

In business applications, web services are proven to be composed in a complex manner [16]-[17]. Moreover, IBM WebSphere could support standardized web services and cooperate with the Microsoft workflow tool. The BEA WebLogic server not only supports web services and XML, but also composes new services. Web services are extendable techniques, especially on developing a large system.

3. System Architecture

Service Oriented Architecture (SOA) is an emerging architectural style. The major ideas of SOA are that service elements are granularly defined and constructed, service interfaces are clearly standardized for composing new services, and the services built by following SOA are reusable. By composing services iteratively, a new system could be formatted for serving a specific domain. A SOA-based system usually includes three key features: software components, services elements and business processes. Web service is one of the most important ways to implement SOA. Web service requires XML-based techniques, such as XML, WSDL, SOAP, and UDDI.

The proposed personalized service recommendation system (PSRS) was built by following the SOA principles. The service elements in PSRS are distributed in different sub-systems. In PSRS, web services are evolved by three generations. First, a number of simple web services are implemented and usually used for query and response. Second, composite web services are derived from the simple ones to form more complex applications. Third, collaborative web services are continually emerging. These dynamic services could automatically support business agility. The architecture of PSRS would be flexible and extendable.

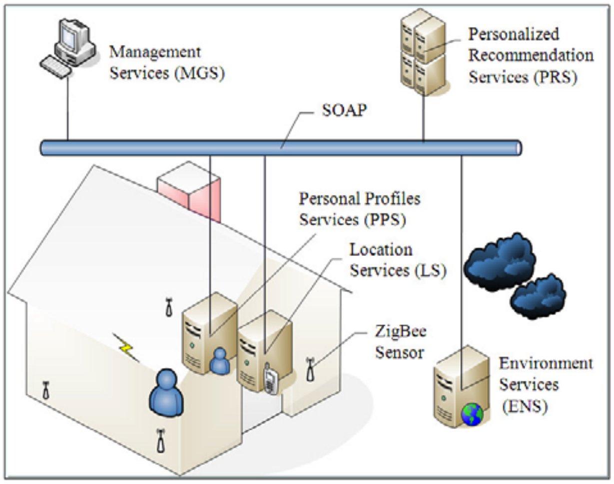


Fig. 1 The distributed Services in PSRS

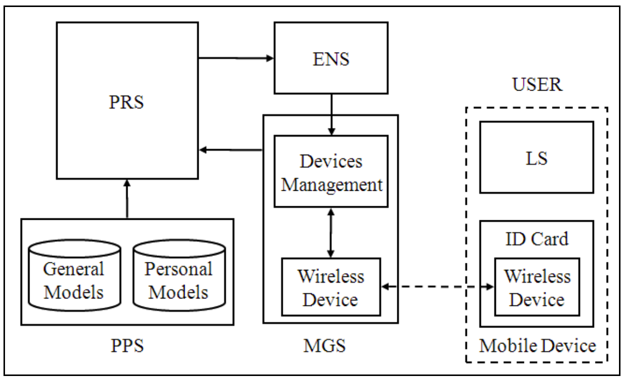


Fig. 2 System Architecture

The services in PSRS are developed according to SOA principles, too. There are three key benefits. By upgrading service components, the system performance could be improved at pace and the faults could be gradually reduced. Second, the system services could be enriched by increasing the number of service components. The system would become progressively better and friendly. Third, users' demands could be fitted in with the dynamic services composition. We could model user's preferences and compose new services for further recommendation. In order to keep the flexibility and extensibility, the services in PSRS are distributed in different service groups. They are explained in the following sub-section and shown in Fig. 1.

3.1 Grouping Services

(1) Personal Profiles Services (PPS): Personal profiles are key factors when recommending services to the person. Data stored in personal profiles could be classified into static data and dynamic data. A person's name, ID, sex, and blood type are categorized as static data. Dynamic data is composed of personal information which is possibly variable, such as age, habit, health status, behavior feature, service level and authority. The dynamic data should be automatically collected and analyzed by the information systems.

(2) Location Services (LS): The person's locations are usually key factors in judging the person's behavior models. For example, when a person moves close to certain facilities, this represents the possibility of use of the facilities. A person who moves from one position to another also represents specific activities, such as entering or exiting a room. Even a person who keeps motionless for a period of time would possibly represent some meanings. Furthermore, moving speed, pattern, and displacement are also key factors for modeling the person's behaviors.

(3) Environment Services (ENS): The environmental services could publish the contexts statuses and provide query services for the other services. Through the ENS, the others services could get the contextual statuses for further recommendable control. For example, the contexts are date, time, temperature, brightness, weather, noise and so on. The environmental devices would also be able to be controlled by ENS since the devices conditions could be simply queried. Furthermore, the services could thus be recommended to automatically control the devices for fitting users' preferences.

(4) Management Services (MGS): The services are responsible for managing the other services and some functionality. The other services would be registered and published in UDDI server and managed by MGS. The user's authority in PSRS would be managed by MGS as well.

(5) Personalized Recommendation Services (PRS): The models of personal activities are analyzed and built by PRS. Personalized services are recommended according to the personalized models, which are tuned by the pre-defined general model and personal behaviors. Personal services in the digital home could be automatically triggered before the user manually controls them. For instance, we could preset control the status of air conditioner, lighting, television, exercise devices among other devices.

For the user's scenarios of this paper, the user takes a mobile measurable device with wireless ID card. The fixed homebox in the house could receive the user's bio-signals and locations from the wireless sensor. At the same time, the MGS could acquire these data and those contexts from the environment. Then, the PSRS could actively select the adaptive services by the previous contexts and the personal models. The PSRS architecture is shown in Fig. 2.

3.2 Generating Recommendable services

The recommendable services are reasoned out by the contexts, personal models and the user's locations in the PSRS. The static factors and rules are pre-defined by the web-based editor and stored in the knowledge base. The dynamic factors used for triggering rules are dispatched from the PPS, LS and ENS. The services would thus be recommended by following the actions of the triggered rules. In addition, the dynamic rules are formulated by the tuned personal models and the factors. If no personal models exist, the general model would be selected. The personal models are carefully described in Section 4. The ruling outputs might be automatically used for passing messages or controlling devices in the digital home. Its outputs could even call on a series of others services. The decision process is shown in Fig. 3.

4. Personalization Models

In PSRS, the primary contexts used to personalize services are the personal models, locations and his/her health statuses. The user could select service modes, such as manually setting devices or automatically recommendation services. If the user enters in the service scope, his/her ID will be sensed. The user's locations will also be identified to trigger recommendable services.

The interoperability of service providing is shown in Fig. 4. If there are existing personal models, the models will be loaded to bind the activity patterns and to select the proper items. The parameters of the devices could be set by the quantitative items. For example, the values of air conditioner and lamplights would be automatically set by following the user's preference. The device usage progress of the user will be recorded to update the personal models, which are built and analyzed through personal modeling.

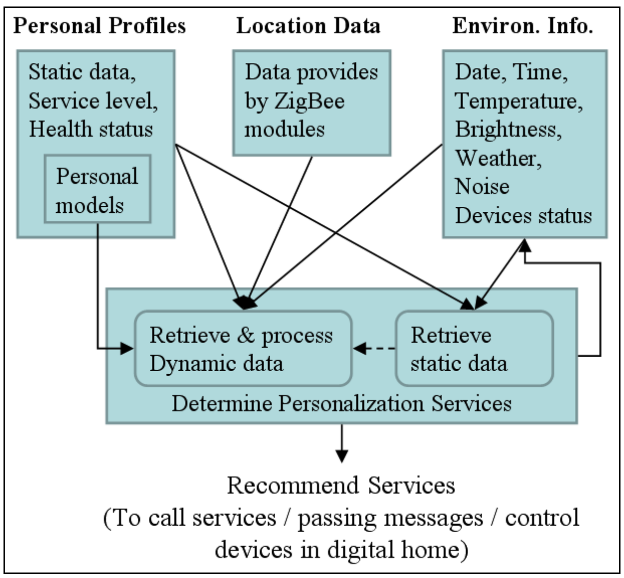


Fig. 3 Recommendable services Generation

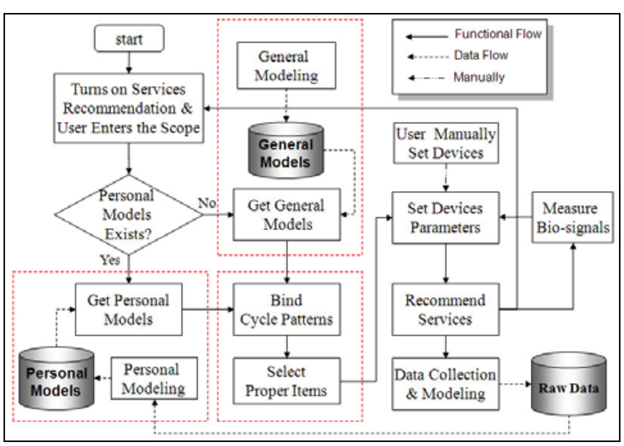


Fig. 4 Interoperability of service providing

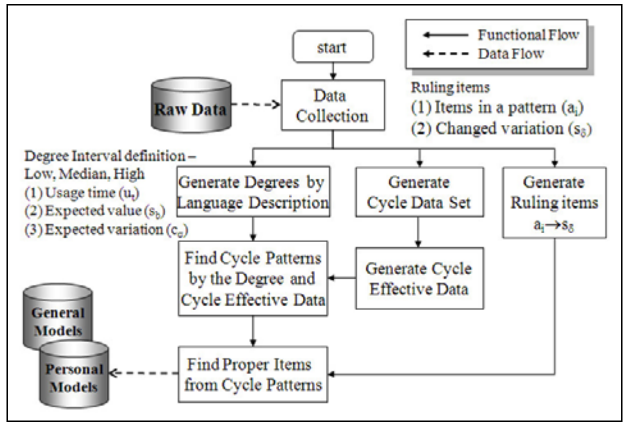


Fig. 5 Personal models Generation

If there are no existing personal models, the general models will be recommended to bind the user's patterns. General parameters will also be set in those devices. If the user manually set the devises during the progress of usage, the user's intension will be recorded for tuning and building new personal models. Personal models of the user would be available for use next time. Undoubtedly, the user could manually set the devices anytime. He/she could manually set the music volume or TV channel for instance.

4.1 Personal Modeling

The user's raw data is collected from recording the interactions between the individual and the devices. Through analyzing the raw data, personal models could be generated. A user's activity patterns will cyclically occur in the same conditions. The cycle effective data are collected by mining the raw data. By discretion the effective data and incorporating them into distinct degrees, the cycle patterns could be found. A cycle pattern is combined with the ruling items, which are mapped to the functions on the devices. Therefore, the cycle patterns could be bound to serve the user. The flow of personal models generation is shown in Fig. 5. The personal models are stored in data repositories. As mentioned in the previous section, personal models could also be updated by the new raw data. Likewise, the general models can be generated by the same process in Fig. 5. However, the differences are that all uses' activities are selected for analyzing.

4.2 Ruling Items and Cycle Patterns

Personal models are stored in a rule-based database. The rules combined with factors and formulas. There are three kinds of factors: 1. Event – dynamic factors; 2. Status – static factors; 3. Compound – composite factors. Each factor has its own identified number. The static factors are represented by negative numbers, while the dynamic factors are represented by positive numbers. The formulas are combined by factors, and are stored with IF-ELSE format in a database. In order to edit the formulas, these are represented as mathematical equations on the website. For instance, the formula could be represented as "2 + 3 + 7 + -8 = 10". The numbers of the formula are factor identifications. The symbol "+" refers to the sequences of the occurred factors. Each launched formula corresponds to an active service. Formulas could be iteratively launched by the dynamic factors during the recommendation. The composite services will be appeared by the iterations.

5. Experiments and Verification

5.1 Experiment Environment

The PSRS was implemented by the C#.NET programming language. We integrated many types of devices to verify our PSRS. A computer server was remotely installed for serving web service techniques. A laptop was included to connect the ZigBee coordinator for receiving the user's locations. The ZigBee modules contain six CC2430 reference nodes, two CC2431 blind nodes and one CC2430 coordinator. The ZigBee modules were purchased from Texas Instruments. One programmable logic controller (PLC) was linked to the laptop by the RS232 interface. The PLC was used for controlling home devices, such as three color lights (red, green, white), one electric fan and one doorbell. The homebox and bio-server are provided by the Institute for Information Industry (III). These facilities are used for measuring the use's bio-signals.

5.2 Test Cases 1 — *The user activity patterns*

The cycle patterns are pre-defined in the web-based editor. These patterns are shown in the following:

(1) Pass through POS1: Turn off green light
(2) Pass through POS2: Turn on green light
(3) Pass through POS3: ring doorbell
(4) Pass through POS5: ring doorbell
(5) Clockwise (POS3+POS2+POS1+POS5+POS3): flash red light
(6) Counter-clockwise (POS1+POS2+POS3+POS5+ POS1): flash white light

A user would launch the services if his/her activities matched the pre-defined patterns. The scenario is shown in Fig. 6. The blue arrows display the counter-clockwise pattern. The yellow arrows display the clockwise pattern. The experimental results showed that the PSRS could correctly execute distinct services based on the user's activity patterns.

5.3 Test Cases 2 — *Multiple users login service*

In this scenario, two users will enter the sensing scope of the homebox. The first user's profile will be pre-loaded into the homebox if he/she enters the scope. He/she could automatically login and then measure his/her bio-signals. If he/she finishes and leaves the sensing scope, he/she would also automatically logout. The second user could then automatically login and use the homebox. The users don't need to manually operate the login process. The login service could automatically work in a multi-user environment. The sequences of the occurred activities are shown in Fig. 7. The yellow arrows are the first user's activities. The blue arrows are the second user's activities. The cycle patterns are shown in the following:

(1) Enter POS4 + none user: ring doorbell
(2) Enter POS4 + not login: automatically login
(3) Login + Leave POS4 = login out

5.4 Test Cases 3 — *Flexible services composition*

For services recommendation, the services could query contexts statuses through web services. In order to show the flexible services composition, the system operator will modify the defined cycle patterns and ruling items. The flexible cycle patterns will then provide different services. The scenarios are shown in Fig. 8. The blue arrows are the first activities. The yellow arrows are the second activities. The green arrows are the third activities. The cycle patterns are shown in the following:

(1) Enter POS5 + Leave POS5: Leave POS5 (Ring doorbell)
(2) Enter POS5 + Evening = Turn on light
(3) Turn on light manually + Leave POS5: Turn off the light automatically

After modifying the cycle pattern:

(4) Enter POS5+Not Evening: Power on a electric fan
(5) Power on the electric fan + Leave POS5: Power off the electric fan

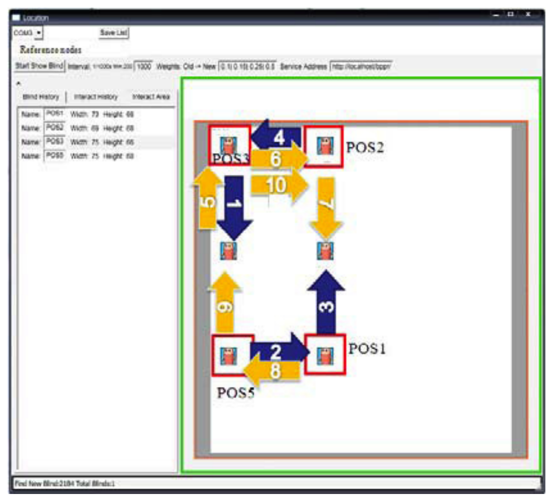


Fig. 6 User Activity Patterns (ZigBee Localization Interface)

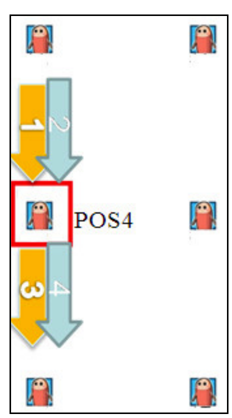


Fig. 7
Multiple Users

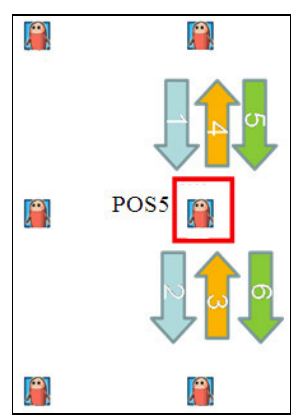


Fig. 8
Flexible Services

6. Conclusion and Future Work

In this paper, we developed a personalized service recommendation system (PSRS) in a home-care environment. The PSRS has capability of providing proper services based on the user's preferences. For the recommendable services, we explored the processes and data sources of generating the recommendable service. Furthermore, we construct personal models to record the user's activities and habits. Through the personal models, the system will be able to automatically launch to safety alert, recommendable services and healthcare services in the house. In the future, the proposed system and models could even carry out the mobile health monitor and promotion in a home-care environment.

7. Acknowledgements

This research was partially supported by the second Applied Information Services Development and Integration project of the Institute for Information Industry (III) and sponsored by MOEA, Taiwan R.O.C.

8. References

[1] Chi-Lu Yang, Yeim-Kuan Chang and Chih-Ping Chu, "Modeling Services to Construct Service-Oriented Healthcare Architecture for Digital Home-Care Business," in Proceedings of the 20th International Conference on Software Engineering and Knowledge Engineering (SEKE'08), pp. 351-356, July, 2008.

[2] B.P.S. Murthi, and Sumit Sarkar, "Role of Management Sciences in Research on Personalization," Management Science, Vol. 49, No. 10, pp. 1344-1362, Oct. 2003.

[3] Asim Ansari and Carl F. Mela , "E-Customization," Journal of Marketing Research, Vol. 40, No. 2, pp. 131-145, May 2003.

[4] Kar Yan Tam, Shuk Ying Ho, "Web Personalization: Is It Effective?," *IT Professional*, Vol. 5, No. 5, pp. 53-57, Oct. 2003.

[5] Hung-Jen Lai, Ting-Peng Liang and Y.-C. Ku, "Customized Internet News Services Based on Customer Profiles," in Proceedings of the 5th international conference on Electronic commerce, pp. 225-229, 2003.

[6] James Pitkow, Hinrich Schütze, Todd Cass, Rob Cooley, Don Turnbull, Andy Edmonds, Eytan Adar and Thomas Breuel, "Personalized Search," Communications of the ACM, Vol. 45 Issue 9, pp. 50-55, 2002.

[7] Josef Fink, Alfred Kobsa and Andreas Nill, "User-Oriented Adaptivity and Adaptability in the AVANTI project," in Conference Designing for the Web: Empirical Studies, Microsoft, Redmond, WA, Oct. 1996.

[8] Peter Brusilovsky, "Methods and techniques of adaptive hypermedia," Journal of User Modeling and User Adapted Interaction, Vol. 6, No. 2-3, pp. 87-129, 1996.

[9] Wolfgang Wahlster and Alfred Kobsa: Stereotypes and User Modeling. Springer. User Models in Dialog Systems, pp. 35-51, Springer, Berlin, Heidelberg, 1989.

[10] Chin, D.N. KNOME: Modeling what the User Knows in UC. User Models in Dialog Systems, pp. 74-107. Springer, Berlin, Heidelberg, 1989.

[11] M. Schneider, Hufschmidt, T. Kühme, and U. Malinowski, "User modeling: Recent Work, Prospects and Hazards," in Adaptive User Interfaces: Principles and Practice, , 1993.

[12] Tim Bray, Jean Paoli, C. M. Sperberg-McQueen, Eve Maler, François Yergeau eds. Extensible Markup Language (XML) 1.0 (Fourth Edition), World Wide Web Consortium (W3C) Recommendation, Nov., 2008.

[13] David B., Canyang Kevin L., Roberto C., Jean-Jacques M., Arthur R., Sanjiva W. et al. Web Services Description Language (WSDL) Version 2.0 Part 1: Core Language. W3C:http://www.w3.org/TR/wsdl20-primer/, 26 June 2007.

[14] Martin G., Marc H., Noah M., Jean-Jacques M., Henrik F., Anish K., Yves L. SOAP Version 1.2 Part 1: Messaging Framework (Second Edition), W3C: http://www.w3.org/TR/2007/REC-soap12-part1-20070427/, 27 April 2007.

[15] Tom B., Luc C. and Steve C. et al., UDDI Version 3.0.2. OASIS: http://uddi.org/pubs/uddi-v3.0.2-20041019.pdf, 19 October 2004.

[16] Brahim Medjahed, Atman Bouguettaya, Ahmed K. Elmagarmid, "Composing Web services on the Semantic Web," The VLDB Journal, 12(4), pp.333-351, Sep. 2003.

[17] W.M.P. van der Aalst, "Don't go with the flow: Web services composition standards exposed," IEEE Intelligent Systems, 2003.

[18] Chi-Lu Yang, Yeim-Kuan Chang and Chih-Ping Chu, "A Gateway Design for Message Passing on SOA Healthcare Platform," in Proceedings of the Fourth IEEE International Symposium on Service-Oriented System Engineering (SOSE 2008) , pp. 178-183, Jhongli, Taiwan, Dec. 2008.

Design and Implementation of OSGi-Based Healthcare Box for Home Users

Bo-Ruei Cao[1], Chun-Kai Chuang[2], Je-Yi Kuo[3], Yaw-Huang Kuo[1], Jang-Pong Hsu[2]
Dept. of Computer Science and Information Engineering[1]
National Cheng Kung University, Tainan, Taiwan
Advance Multimedia Internet Technology Inc., Taiwan.[2]
Institute of Information Industrials, Taiwan[3]
kuoyh@ismp.csie.ncku.edu.tw

Abstract

With the development of Internet technology and computing costs continue to decline, the realization of being the dream of digital home life. In this paper, we will explore Open Service Gateway initiative (OSGi) technology to develop a transferable framework to support a cross-platform environment of health-care service Intelligent Home Health-Care Box platform to achieve the following objectives: (1) To develop remote physiological signal measurements (2) To take advantage of OSGi (Open Service Gateway initiative) to construct a transferable framework for embedded computing. (3) To reduce the program size in an embedded system and upgrade the performance in run time stage. In this constructed environment, home user can demand services prompted by a service-discovery mechanism and interact with health-care devices through network. In other words, networking, intelligence and multimedia are the guidelines to investigate and develop the residential information system who will serve people at home in a friendly manner and improve the quality of home living.

Keywords: *Embedded System, medical equipment, Home Health Care, OSGi, Remote Health Care.*

1. Introduction

As medical science and technology developing, the average age of humans is also growing and the social structure is aging. According to the definition by World Health Organization (WHO), an ageing country is the one whose 7% population has the age greater than 65 years old. The WHO estimates that, in 2020, most developed countries will encounter the problem of ageing population. In particular, Japan, North Europe and West Europe will have ageing population more than 20%.

In Taiwan, Taiwan indeed becomes an aging country in September 1993. At present, Taiwan's population aged over 65 have more than two million. According to the latest survey, the aged population has exceeded 9.1% in Taiwan. Council for Economic Planning and Development (CEPD) in Taiwan estimates that, in 2031, the 65-year-old population will reach 19.7 percent of the total population. In the other words, every five people have one aged people in that time. The degradation of physiological functions makes these aged living alone occurred inconvenient situation, especially in the aged suffering chronic diseases. Therefore the health-care issues of the aged people have become significant and attractive for research. For aged people, the offering of Long-Distance Home-Care for the aged people with chronic disease such as hypertension will effectively improve the living quality and reduce the burden of hospital-care system. Long-Distance Home-Care provides such functions as personal emergency rescue, long term physiological signal monitoring (an aspect that is keenly important to physiological monitoring equipment industry) that uses electronic blood pressure devices, blood glucose meters, and clinical thermometers. This Distance-Care approach will improve incomes and services for hospital systems, telecommunication companies, and security companies. It is projected that production value of Long-Distance Home-Care in Taiwan is expected to reach the scale of NT$7 billions by 2010. [1]

Not only in Taiwan, has research showed that the global home-care market is growing rapidly by 20% each year. In 2006, the scale of the global home-care market was about USD 71.9 billions. It is predicted that it will increase to USD 79.6 billions in 2010. If all related industry and institutional services were included then that market scale would be even larger. What these statistics underscore is the need for home and institutional care for senior citizens to let them play a greater part in modern day society.

The Intelligent Home Health-Care Box platform [2] already has been achieved using network technology and information technology to provide an intelligent assisted care system. Allow remote monitoring by the medical staff to obtain measurements of physiological signals of patients, will be able to greatly improve the current blind spot in the regular care visits. Acquired physiological signals are more immediate and can also reduce the time to measurement to the home. In addition, measurement of physiological signals has become medical records of patients. Therefore the Intelligent Home Health-Care Box really can assist in monitoring by the health status of caregivers allows patients to get the best home care environment. But with the development of Internet technology and computing costs continue to decline, the requirement for digital medical care would be growing and diversified. This delivery system will inevitably have to rely on the transfer type can be embedded platform technology to provide remote computing service composition, remote service delivery functions.

In this paper, we will research for transferable framework for embedded computing applies to digital care services. We will base on embedded platform to design a transferable computing technology, this computing technology necessary includes: (1) Construction services provide real-time execution environment. (2) Automated service management and scheduling. (3) The services required for real-time transmission of content. (4) The development of service programs to be able to transfer between different platforms.

In other words, we must construct a model-oriented service structure, to achieve service description, construction services, authentication service and service delivery functions of the target. And use home-care as an example, to carry out situational analysis and build the prototype system developed to validate the technology. We call the system named model-oriented nursing system (MON).

The rest of this paper is organized as follows. The related works of information technology applied on remote nursing application are reviewed in Section 2. In Section3, the proposed remote MON system is presented in terms of architecture, functions and implementation. The Numeric Results of MON system are demonstrated in Section 4. Finally, this work is concluded in Section 5.

2. Related Work

The nursing problem of aged people is a critical issue in most developed countries, such as United State, Japan and Europe. In US, the care demand of aged people facilitates the market growth of home-nursing services. And home-nursing services gradually become a demand in trend and attract lots of researches. The previous research works mainly focus on how to employee the modem information and networking technologies to establish computer-aided home nursing systems. For example, Wong et. al. [3] proposed a lifestyle monitoring system (LMS) using passive infrared movement detector (PIR) to detect the behavior and body temperature of the cared patient in room. When unusual conditions are sensed by a control box, the control box will deliver the collected data to laboratory for further analysis. N. Noury et. al. [4] proposed a fall sensor composed of infrared position sensors and magnetic switches to remote monitoring human behavior. Once the monitored person is falling down, the fall sensor notifies the remote care center through RF wireless networks. The care center will assign neighboring rescuer to deliver the in-time treatment to save that person.

K. Doughty *et. al.* [5] presented a dementia patients living alone monitoring system to monitor the daily behaviors of dementia patients and to generalize an on-going dementia lifestyle index (DLI) for each patient. The DLI is empirically useful to verify the effectiveness of the medical treatment and to guide the treatment of each patient. American TeleCare, Inc. [6] established in 1993 has 9500 market sharing of home telemedicine products including a patient station connecting Central Station by a phone line to transport the signals of telephonic stethoscope, blood pressure meter and oximeter. Patient station monitors the patient status and delivers the data to Central Station. However, patient station does not analyze the collect signals and response to the exception. Nigel H. Lovellv et. al. [7] demonstrated a web-based approach to acquisition, storage, and retrieval of biomedical signals. The home patient monitored by a terminal to record his blood pressure, breath, pulse. The records are delivered to and stored at hospital database. Clinic doctors will heal the patient with more useful medical information.

Most patient-monitoring applications [8-12] do not allow remote access control from care center. This kind of system can not be managed remotely. The proposed MON provides remote access control function to manage. By this way, the wide-deployed MON are feasible to maintain and the maintain cost can be significantly reduced. The remote access control is designed based on Open Service Gateway initiative (OSGi).

Besides, since MON is operated across Internet, the impact of network performance and the requirement of network resource should be studied. Horng et. al [13-14] proposed a delay-control approach to guarantee the quality-of-service (QoS) for home users and a fine granularity service level agreement (SLA) to manage network resource. Huang et al. [15] presented a residential gateway to translate communication protocols, coordinates information sharing and serves as a gateway to external networks for integrated services. The evolving techniques are greatly beneficial for users in home environments. Thus, in this paper, the characteristic of network resource usage caused by the proposed MON is also investigated deeply.

3. OSGi-based Healthcare Homebox of Model-Oriented Nursing System

The system architecture of the MON is depicted in Fig.2. There are three parts to introduce: Hardware platform (Intel Xscale 270-S), System Software and Functional module. Hardware platform is developed on Intel Xscale 270-S. This platform includes processor, flash, sdram and many interfaces. The processor PXA270 [16] is designed to meet the growing demands of a new generation of leading-edge embedded products. Featuring advanced technologies that offer high performance, flexibility and robust functionality, the Intel PXA270 processor is packaged specifically for the embedded market and is ideal for the low-power framework of battery-powered devices. The MON platform can use external 5V power supply to work or use built-in 3500mA/h lithium battery to work. The battery can supply power more than 5 hours and the platform support power supply charge or USB charge. Therefore it is very suitable for mobile devices. The main specifications are as follows Intel Xscale 270-S hardware specifications, finishing in the table 1.

System software has four parts: (1) Hardware driver (2) Operation system (3) Embedded JVM (4) OSGi framework

(1) Device driver:

The Device driver contains RS232, Ethernet, Frame buffer, Touch panel and Sound. RS232 driver for connect RS232 medical instruments and collect the measurement from instruments. Ethernet driver is for remote monitoring. Frame buffer driver is for display, Touch panel driver is for user control. Sound driver is for alert.

(2) Operation system:

We use Linux, Kernel version 2.6.9 for our Operation system. The library we use uClibc because it is most suitable embedded Linux.

(3) Embedded JVM:

At present, an emerging issue in JVM is that applies JVM to use on embedded system. Java standards are dominated by the Sun Company. Sun's JVM and other Java API have been regarded as the standard Java platform. Any implementation of Java platform must be compatibility with Sun's JVM platform as a top priority. But a long time, Sun has been reluctant to put own Java platform open. When use java may be a concern about authorized, indirectly hinder the promotion of Java. But that has inspired many of the JVM have Open source, such as Kaffe, Jikes RVM, JamVM and so on. In the last year, Sun donated the Java technology to Open Source Community (OpenJDK [17]). But OpenJDK still have the problem that it rarely platforms supported. Because we need a JVM can support many platforms, we still need other Open Source issue.

Existing implementation of Open Source Java usually based on GNU Classpath [18]. GNU Classpath 1.0 will be fully compatible with the 1.1 and 1.2 API specifications, in addition to having significant (>95%) compatibility with the 1.3, 1.4, 1.5 and 1.6 APIs. As a result of Classpath have significant compatible with JAVA API, many Implementation of Open Source JVM use Classpath to its API. And in Open Source JVM, we choose JamVM [19] as our platform's JVM

JamVM is a new Java Virtual Machine which conforms to the JVM specification version 2 (blue book). In comparison with most other VM's (free and commercial) it is extremely small. JamVM's interpreter is highly optimized, incorporating many state-of-the-art techniques such as stack-caching and direct-threading. The stack-caching is keeping a constant number of items in registers is simple, but causes unnecessary operand loads and stores. E.g., an instruction taking one item from the stack and producing no item (e.g., a conditional branch) has to load an item from the stack, that will not be used if the next instruction pushes a value on the stack (e.g., a literal). It would be better to keep a varying number of items in registers, on an on-demand basis, like a cache. The direct-threading is used in order to save memory space, the compiler will generate the corresponding subroutine native code, but in an indirect line of the structure of serial code, the compiler does not directly generate native code, and are independent of each subroutine in the library, and at compile time, resulting in subroutine inside the library of memory addresses, and then to implementation, the overhead line through a series of code memory addresses, complete executive action. As most of the code is written in C JamVM is easy to port to new architectures. So far, JamVM supports and has been tested on the following OS/Architectures, including PowerPC, PowerPC64, i386, ARM, AMD64, i386 with Solaris/OpenSolaris.

In addition, JamVM is designed to use the GNU Classpath Java class library. A number of classes are reference classes which must be modified for a particular VM. These are provided along with JamVM. JamVM should always work with the latest development snapshot of Classpath.

(4) OSGi framework:

OSGi Framework implements a complete and dynamic component model, something that does not exist in standalone Java/VM environments. Applications or components (coming in the form of bundles for deployment) can be remotely installed, started, stopped, updated and uninstalled without requiring a reboot; management of Java.

The OSGi Alliance [20] (formerly known as the Open Services Gateway initiative, now an obsolete name) is an open standards organization founded in March 1999. The Alliance and its members have specified a Java-based service platform that can be remotely managed. The core part of the specifications is a framework that defines an application life cycle management model, a service registry, an Execution environment and Modules. Based on this framework, a large number of OSGi Layers, APIs, and Services have been defined.

We choose OSCAR (Open Service Container Architecture) [21] as our platform's OSGi framework because it is a tiny OSGi framework. The program size of OSCAR OSGi framework only 388 KB in run time, other OSGi framework such as Knopflerfish and Equinox all need more than 5 MB program size in run time. At present, OSCAR has been renamed as felix [22], being developed by the Apache.

In the OSGi framework, the software can Independent implement the function completely known as the Bundle. In terms of Implementation, Bundles are normal jar components with extra manifest headers. A Bundle object is the access point to define the lifecycle of an installed bundle. The lifecycle of bundle is show in Fig.3.

Each bundle installed in the OSGi environment must have an associated Bundle object. A bundle must have a unique identity, a long, chosen by the Framework. This identity must not change during the lifecycle of a bundle, even when the bundle is updated. Uninstalling and then reinstalling the bundle must create a new unique identity. A bundle can be in one of six states: UNINSTALLED, INSTALLED, RESOLVED, STARTING, STOPPING, and ACTIVE. Values assigned to these states have no specified ordering; they represent bit values that may be ORed together to determine if a bundle is in one of the valid states. A bundle should only execute code when its state is one of STARTING, ACTIVE, or STOPPING. An UNINSTALLED bundle can not be set to another state; it is a zombie and can only be reached because references are kept somewhere. The Framework is the only entity that is allowed to create Bundle objects, and these objects are only valid within the Framework that created them.

The main purpose of OSGi standard is to provide a complete point-to-point service delivery solution between remote care center and local MON platforms. Therefore, the OSGi defines an open platform for user can download applications from remote care center and install and execute automatically in any time. We hope that through this open platform, developed by different vendors of software and equipment services can communicate and use with each other. Functional module has User Interface and four bundles, the four bundles are mapping three scenario. The mapping schematic was displayed in Fig 4. There are three scenarios developed in this work as follows.

Scenario 1: Remote monitoring

Remote user can access the physiological information record in MON platform or Get the emergency message from Alert Key Bundle through Web Server Bundle. In addition, User also can use Web Server Bundle to remote install, start, stop, upgrade bundle.

Scenario 2: Emergency call

When emergency, user can press the emergency button on the User Interface (in the title button). In this time, Alert Key Bundle will be start and send out an emergency signal through Web Server Bundle to notify remote care center.

Scenario 3: Physiological signal measurement

The MON platform has a touch panel can let user touch screen to control the connected RS232 physiological signal measuring apparatus such as ventilator, blood pressure monitor and pulsimeter through GUI bundle. And use RS232 Interface Bundle to change the physiological signal from RS232 physiological signal measuring apparatus and record that in to MON platform. Even let this information upload to remote care center for monitor and record.

4. Numeric Results

There front view the hardware platform, Intel Xscale 270 to develop healthcare homebox is shown in Fig. 5. There are a mother board and a TFT-LCD in this Xscale platform. Based on this platform, the developed system software, application software and user-interface software are ported and integrated to realize the three scenarios as described as mentioned. The front-view of the user-interface is depicted in Fig. 6. Instead of keyboard, a GUI interface with touch panel is employed for users. Such a friendly design is more valuable and feasible to home users.

Certainly, the system performance is also evaluated to verify the improvement of java virtual machine. There are two key performance index (KPI) chosen to evaluate the performance, including starting time and memory utilization of JVMs. Three kinds of JVM technology are compared. They are Kaffe, JamVM and embedded J2SE. The comparison results are shown in Fig. 7. Obviously, the adopted JamVM demonstrates the better execution than Kaffe. The performance of JamVM is quietly close to typical embedded J2SE. Although JamVM and Kaffe all run JAVA program as interpreter, JamVM's interpreter is highly optimized. In Section3, we talk about the JamVM incorporating many state-of-the-art techniques such as stack-caching and direct-threading. These techniques let the JamVM has high performance interpreter so it can start OSCAR and load OSCAR bundle profile quickly.

However, the program size of embedded J2SE is larger than the JamVM. As shown in Table 2. JamVM's program size is only 15.2 in run time. JamVM and Kaffe all use the Classpath to its JAVA library. So JamVM's program size is close to Kaffe.

5. Conclusions

In this paper, we propose a model-oriented nursing system, called MON system to enactive home health care system and to satisfy the requirements for the next generation of home health care system. MON system cooperating with remote care centers plays an important role to realize a smart home with health-care applications. Through MON system, patient enjoys medical information services and on-line interaction with staffs in care center. Care center has a continuous monitoring of medical measurements for each home patient. The experimental results depict that MON effectively enhances the nursing quality of home patient through information and networking technologies. Besides, the performances of deployment are also evaluated. The interaction between the patient and the service center is the key advantage of the proposed system and also is the trend. The proposed MON demonstrates a feasible approach to enhance the home healthcare service to meet the requirements of the aged people and the coming ageing society. In particular, this MON platform achieve remote operation, maintain and administration (OAM) based on OSGi Standard, such as including software module (bundle) remote install, update and control. This is an innovation and revolution feature for remote health care. And this feature make remote health care more flexible and immediately.

Acknowledgment

This paper is based partially on work supported by the National Science Council (NSC) of Taiwan, R.O.C., under grant No. NSC97-2218-E-006-014 and Institute for Information Industry of Taiwan, R.O.C..

References

[1] A SenCARE & industry backgrounder - 2009SenCARE, http://www.sencare.com.tw/presscenter/news_view.shtml?docno=4842
[2] M. F. Horng *et. al* ,"Development of Intelligent Home Health-Care Box Connecting Medical Equipments and Its Service Platform," Proc. of IEEE 9th International Conference on Advanced Communication Technology, (ICACT2007), CR-ROM, Korea, 2007.
[3] C. Wong and K. L. Chan, "Development of a portable multi-functional patient monitor," Proc. of the 22th Annual EMBS Int'l Conf. vol. 4, pp.2611-2614, 2000.
[4] N. Noury *et al.,* "Monitoring behavior in home using a smart fall sensor and position sensors," Proc. of the 1st Annual International Conference On Microtechnologies in Medicine and Biology, (2000) 160-164
[5] K. Doughty, "DIANA-a telecare system for supporting dementia sufferers in the community," Proc. of the 20th Annual EMBS Int'l Conf. vol. 4, (1998) 1980-1983
[6] AmericanTeleCare, http://www.americantelecare.com/
[7] N. Lovel, et al.,"Web-based Acquisition, Storage, and Retrieval of Biomedical Signals", IEEE Engineering in Medicine and Biology Magazine, vol. 20, no. 3, (2001) 38-44
[8] Kevin Hung and Yuan-Ting Zhang, "Implementation of a WAP-based telemedicine system for patient monitoring," IEEE Transaction on Information Technology in Biomedicine, vol. 7, no. 2, (2003) 101-107
[9] F. Magrabi, N. H. Lovell, and B. G. Celler, "Web based longitudinal ECG monitoring," Proc. 20th Annu. Int. Conf. IEEE EMBS, vol. 20, no.3, (1998) 1155-1158
[10] S. Park et al., "Real-time monitoring of patient on remote sites," Proc. 20th Annu. Int. Conf. IEEE EMBS, vol. 20, no. 3, (1998) 1321-1325
[11] B. Yang, S. Rhee, and H. H. Asada, "A twenty-four hour tele-nursing system using a ring sensor," Proc. 1998 IEEE Int. Conf. Robotics Automation, (1998) 387-392
[12] Yonghong Zhang, Jing Bai and Wen Lingfeng, "Development of a home ECG and blood pressure telemonitoring center," Proc. 22 Annu. Int. Conf. IEEE EMBS, (2000) 23-28
[13] Horng, Mong-Fong and Kuo, Yau-Hwang," A rate control scheme to support isochronous delivery in wireless CDMA link by using state feedback technique," Proc. of IEEE 6th International Conference on Advanced Communication Technology, 2004 (ICACT2004), vol. 1, pp. 361-366, Korea, Jan. 2004.
[14] Chien-Chung Su, Wei-Nung Lee, Mong-Fong, Horng, Jeng-Pong Hsu and Yau-Hwang Kuo, "Service Level Agreement: A New Bandwidth Guarantee of Flow-level Granularity in Internet VPN, " Proceeding of IEEE 6th International Conference on Advanced Communication Technology (ICACT 2005), vol. 1, pp.324- 329, Korea, Feb. 2005.
[15] W. S. Hwang and P. C. Tseng, 2005, "A QoS-aware Residential Gateway with Bandwidth Management," IEEE Transactions on Consumer Electronics, vol. 51, no.3, pp. 840-848 Aug. 2005
[16] Intel PXA270 processor, http://www.intel.com
[17] OpenJDK, http://openjdk.java.net/
[18] GNU Classpath, http://www.gnu.org/
[19] JamVM, http://jamvm.sourceforge.net/
[20] OSGi Alliance, http://www.osgi.org/
[21] OSCAR, http://forge.ow2.org/projects/oscar/
[22] Felix, http://felix.apache.org/site/index.html

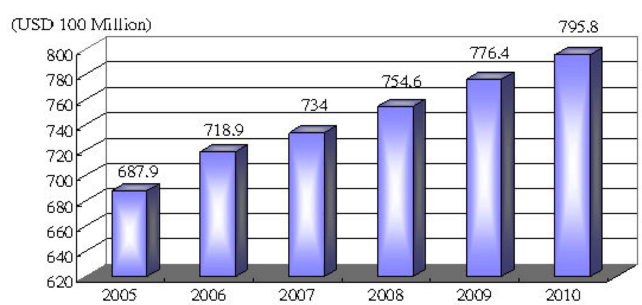


Fig.1 Market Scale of the Homecare Industry
Source: Department of Industrial Technology, MOEA, 2008/04

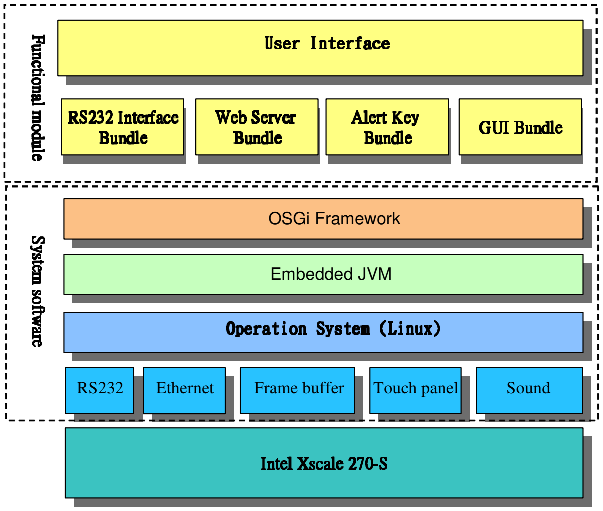


Fig. 2 System architecture of MON

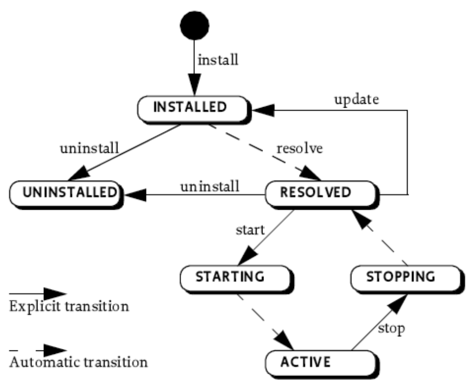


Fig. 3 The lifecycle of bundle

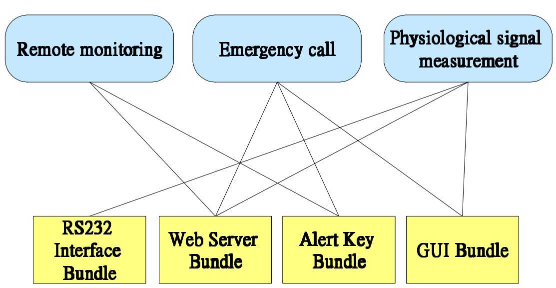


Fig. 4 The scenario and bundles mapping schematic

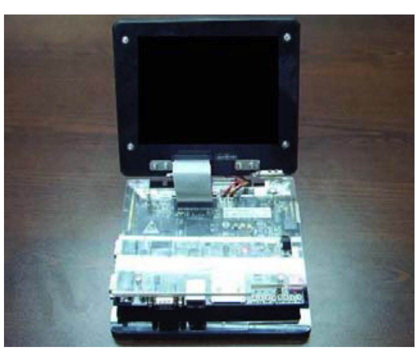


Fig. 5 MON platform

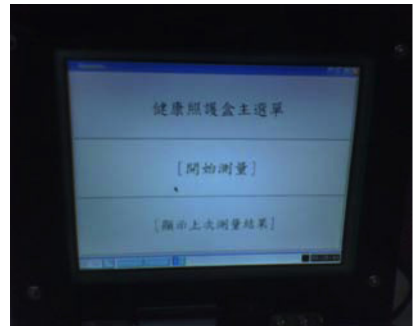


Fig. 6 User interface appearance

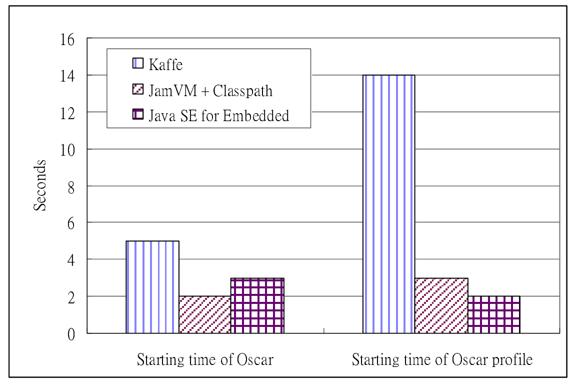


Fig. 7 Performance comparison of various JVM modules in Healthcare home box.

Table 1 Intel Xscale 270-S specifications

CPU	Intel Xscale PXA270 520MHz
SDRAM	64MBYTE
FLASH	32MBYTE
Interface	♦ LCD Monitor: Sharp 3.5 "TFT 320 * 240 ♦ Touch panel: 3.5 "four-wire touch LCD, UCB1400BE control ♦ Serial: 2 RS232 interface, 1 full-function serial ♦ USB HOST: 1 Host interface ♦ USB CLIENT: 1 Client interface ♦ LED Lamp: 8 LED lights

Table 2 Measurements of system performance

	Kaffe	JamVM	Embedded J2SE
Program Size (MB)	18.5	15.2	25.1

An Approach for Tagging 3D Worlds for the Net

Fabio Pittarello
Università Ca' Foscari di Venezia
Dipartimento di Informatica
Via Torino 155, 30172 Mestre (VE), Italy
pitt@dsi.unive.it

***Abstract*—Free-tagging is one of the leading mechanisms characterizing the so-called web 2.0, enabling users to define collaboratively the meaning of web data for improving their findability. Tagging is applied in several forms to hypermedia data, that represent the largest part of the information on the web.**

In spite of that, there is a growing part of web data made of 3D vectors representing real objects such as trees, houses and people that lacks any semantic definition. Such situation prevents any advanced use of the data contained inside these 3D worlds, including seeking, filtering and manipulation of the objects represented by vectors.

This work proposes a bottom-up approach for adding semantics to such data, based on the collaborative effort of users navigating the web. The paper, after describing the similarities and the differences that characterize tagging for hypermedia and interactive 3D worlds, discusses the design choices that have guided the definition of a specification for inserting tags in the latter environments.

The approach permits the annotation of most 3D worlds compliant with X3D, the ISO standard for describing 3D interactive worlds for the net.

I. Introduction

Because of the availability of faster graphics cards and broader communication networks, the number of 3D worlds for the net is gradually increasing. The application domains are different, ranging from urban studies and tourism to social networking. In most cases the modeling of 3D environments and objects is based on low-level geometric elements like polygonal meshes or, for the most advanced environments, on objects belonging to the family of NURBS surfaces. The authors of 3D worlds implicitly associate a semantics that is recognized by the visitors of the 3D environments; a successful outcome of this process is granted both by the skill of the author and by the existence of a common cultural background shared between the author and the visitor of the 3D world.

Unfortunately, no high-level information related to the objects represented by the polygonal meshes or to their relations is usually available in the files where the 3D information is stored.

The lack of any high-level annotation for the components of these environments prevents any use different from the direct visualization and interaction with the single 3D world. A range of possible interesting uses of such information includes:

- indexing of high-level information by search engines; such information may then be used for seeking different 3D worlds, basing the process on the indexed labels;
- comparison of different 3D world based on the analysis of the similarity of labels;
- automatic presentation of high-level information to the users navigating the 3D environment, associated to the location and to the objects they are currently browsing;
- extraction of semantic objects for examination or for automatically creation of high-level repositories (e.g., a repository of trees extracted from different 3D worlds).

In the last few years there have been a number of proposals for adding semantic information to 3D worlds.

Most proposals are characterized by a top-down approach: low level geometric objects are associated to instances of high-level classes, belonging to predefined domain ontologies (e.g., the *kitchen wall*, belonging to the class *wall*). In these proposals the annotation process is constrained because the user may use only one of the available classes (e.g., the class *wall*) and relations (e.g., the *containment* relation).

This work proposes a different complementary approach based on the free selection and annotation of geometric objects. While this process, widely diffused in the hypertextual web and defined as *tagging*, is characterized by informational fuzziness, it gives a powerful opportunity of labelling objects according to different points of views and lets the high-level semantics of the tagged objects gradually and dynamically emerge.

Altough this work shares with the hypermedia tagging the general concepts and practices, there are some differences and additional issues deriving from the specific application to 3D worlds.

In particular:

- the information objects available for tagging are not clearly identified from the start, as it happens for hypermedia tagging;
- the 3D scene may be populated by vectors characterized by different levels of granularity that can't always be associated to a specific high-level meaning; therefore they may require a preliminary grouping operation before assigning a tag to them; the different grouping choices that might be operated by the users represent an additional variable that adds a level of complexity to the tagging process;
- altough some standard information structures for presenting and navigating the result of the tagging activity may be derived from hypermedia (e.g., the so-called *tag cloud*), 3D worlds may benefit from different presentation

techniques, for avoiding presentation clutter; for example, tags to present may be filtered according to the current position and orientation of the user avatar inside the 3D world.

The rest of the work is organized as follows: Section 2 will consider related works, with a particular reference to the semantic description of 3D environments and to hypermedia tagging; Section 3 will compare tagging for hypermedia and interactive 3D worlds; Section 4 will describe the goals and the design choices of this proposal; Section 5 will show how tags may be included in a standard X3D file for describing objects and spaces; Section 6 will conclude the paper, giving some hints for future development.

II. Related Work

Research related to the semantic annotation of multimedia documents has become increasingly important in the last few years. In the context of the audio-video domain, the Moving Picture Experts Group (MPEG) [1] has defined a set of standards for coding and describing such data. The most interesting standards in relation to this work are MPEG-4 [2] and MPEG-7 [3] [4]. The first specification defines a multimedia document as the sum of different objects and includes an XML based format containing a subset of X3D [5], the ISO standard for describing 3D worlds for the web. The latter specification permits to describe multimedia content of different nature (e.g., MPEG-4, SVG, etc.).

Some interesting proposals [6] [7] [8] use the MPEG-7 standard for annotating the semantics of a 3D scene. Halabala [7] uses MPEG-7 to store scene-dependant semantic graphs related to a 3D environment. Also Mansouri [8] uses MPEG-7 for describing the semantics of virtual worlds. The feature is introduced for enhancing queries and navigation inside 3D environments (e.g., the system can return virtual worlds after semantic queries such as *I am looking for a big chair*).

Concerning the web, the World Wide Web Consortium promotes the definition of a set of languages, rules and tools for high-level description of information. The semantic web is composed by different layers, where the lower one is occupied by the data themselves (expressed in XML) and the higher ones describe - through the introduction of languages such as RDF (Resource Description Framework) [9] and OWL (Web Ontology Language) [10] - the semantic properties of such data. Pittarello et al. [11] propose to integrate such languages in a scene-independent approach for annotating 3D scenes. In this approach the X3D language is used for describing the geometric properties of 3D environments and their associations with high-level semantics, while RDF and OWL are used for defining the scene-independent domain ontology.

The annotation process proposed in [11] includes not only the geometric objects defined into the scene, but also the spaces generated by these objects and inhabited by (virtual) humans. The approach stems from a previous research work [12] aimed at labelling in a multimodal way the environment spatial locations, in order to enhance the user orientation and navigation inside of them.

The technique of free tagging, typical of the so-called web 2.0, permits to final users to annotate documents, giving birth to new structures for organizing information. While these structures, called folksonomies [13], suffer from drawbacks such as homonymy, synonymy and polysemy that are endemic to the bottom-up bulding process, they offer an additional opportunity to label information, standing from the user point of view.

If, according to Boiko [14], content can be defined as the sum of data and associated metadata, we may say that the application of user-specific metadata to data generates multiple contents, derived from the interpretation of data given by different users.

The bottom-up approach is opposed to the classic top-down approach in which a designer defines the information structure of the site [15]. Both approaches have specific points of strength and suffer from drawbacks. That is the reason why some authors have proposed different forms of integration, for composing the need for rigorous classification, increased expressivity and improved findability. While some of the experiences reported in literature are targeted at deriving ontologies from folksonomies [16] [17], other approaches go towards the integration of top-down and bottom-up structures, originating two complementary systems for navigating information [18].

III. Tagging Hypermedia and 3D Worlds: a Comparison

This paper, inspired by the work done in the hypermedia domain, suggests to use tagging as a means to let content emerge from the raw 3D data. As stated in the beginning of this work, the application of semantic labels is particularly relevant in a situation characterized - in the most part of the cases - by the lack of any high-level information.

The lack of this information prevents any use of the data different from what has been conceived by the world author. In most cases the use is related to the simple visualization or interaction inside a specific 3D world.

This situation represents a serious drawback, if compared with what happens in the hypermedia web, where the searching and navigation possibilities rely not only on the structures designed by the information architects of the specific sites, but also on the indexing activity of web crawlers and on the classification activity made by users through tagging.

The possibility to search and browse across a network of different web sites is one of the peculiar features of the web and one of the reasons of its successful affirmation. In contrast, most of the 3D worlds available on the net are separate islands that can't be cross searched, filtered or compared.

Tagging may represent an opportunity for letting information emerge from the raw representation and for building powerful cross-world searching and navigation systems.

This work suggests to use tagging, applied to 3D worlds, for all those situations where an ontology for specific domains is not available or where the existing ontology may be profitably used only by skilled users. For example, an ontology targeted at classic architecture may be profitably used only by subjects

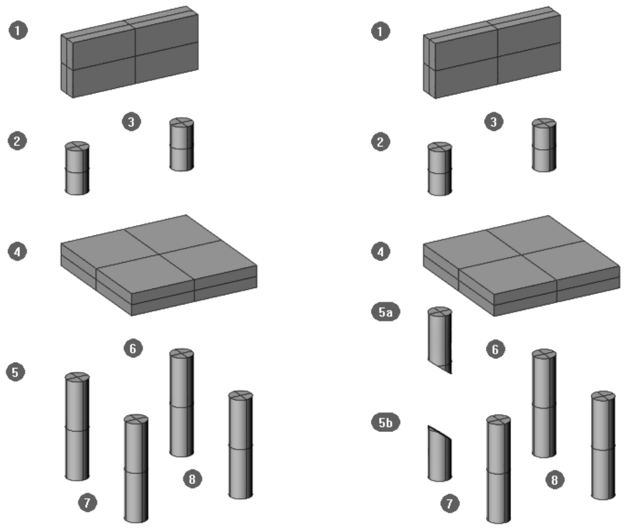


Fig. 1. Different styles for defining the components of a chair

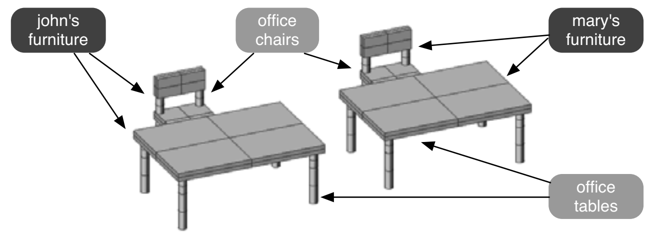


Fig. 2. Different styles for grouping and tagging objects

that are aware of the meaning of terms such as *capital*, *triglyp* or *entablature*. Users that are not trained in the architecture domain might be unable to use such technical terms and they might still want to classify the available information with their own words.

The goal of the proposal is to preserve the same freedom of tagging that is typical of hypermedia tagging systems. For reaching such goal, there are a number of difficulties that are typical of 3D worlds. For hypermedia, the class of objects that may be tagged (e.g., a web page, a video or a photograph) is clearly identified during the design phase and all the tags defined by users will be associated to instances of this class. Besides, during the tagging phase, the targets can be clearly identified. Users don't have to select them among other types of objects, but just specify tags.

That is not true for 3D worlds, where the raw data represent objects belonging to different classes and have different levels of granularity. The modeling process may lead, for example, to use a single mesh for representing a chair or - alternatively - to use different meshes for defining the legs and the seat. Meshes, during the modeling phase, may be grouped, depending on the author habits.

Additionally, the 3D modeling practices may lead to create geometrical objects that don't have a semantics, if considered separately. Fig.1 displays the object *chair* modeled with a different number of components. For what concerns the model on the left, all the components have a semantics that can be easily identified (e.g., the legs, the seat, etc.). The model on the right is characterized by two components, labeled with *5a* and *5b*, that don't have a specific semantics, being only subsets of a leg. Generally speaking, this situation may derive from the fact that a specific set of meshes has been modeled only for obtaining a result in terms of visual presentation rather than keeping in mind the direct association with an high-level meaning. Besides, in some situations, meshes may derive also from some automatic process (e.g., a 3D scanning operation) that may not take into account the object semantics. In such cases it may be necessary to group or split objects, as a preliminary operation for associating a meaning to them.

IV. Goals and Design Choices

One of the goals of our proposal is to give the user the possibility to apply the labels that identify the semantic properties of the objects with the maximum freedom. As stated before, the 3D domain is characterized by different classes of objects that may be tagged and by different levels of granularity. We decided to treat this situation as an additional opportunity for tagging. According to this choice, in our proposal all the geometric objects belonging to the 3D world are taggable. Users are also enabled to define new groups of objects and associate tags to them (e.g., the user may decide to tag the single components of the chair defined in Fig.1 and then to define a group where to put all the components and tag it as *chair*).

Of course we are conscious that different users may decide to define and tag overlapping groups of objects, as can be seen in Fig.2. In this example two users apply different styles for tagging the objects of a room. The first user applies the tags *chairs* and *tables* - evidenced in light gray - after grouping the objects belonging to the same category of furniture; the latter one groups and tags the objects in relation to the owner (i.e., *john's furniture* and *mary's furniture*).

Different tagging styles may represent an issue for the progressive building of the world semantics, introducing a significant amount of informational noise.

On the other side, the opposite choice of forbidding overlapping groups may support informational convergence, but may present additional problems. For example, the association of tags to groups of objects may be restricted to groups of the scene graph or defined by previous users, forbidding the creation of groups that use only a part of the components of existing groups. Unfortunately, following this methodology, inaccurate grouping choices made by previous users can't be further modified. The process may push the tagging activity towards the wrong direction, originating bad semantic associations, such as groups lacking a part of the semantically relevant components.

Some techniques, such as the simple suggestion of the groups already defined by other users, may be an acceptable compromise for reducing the informational noise and support-

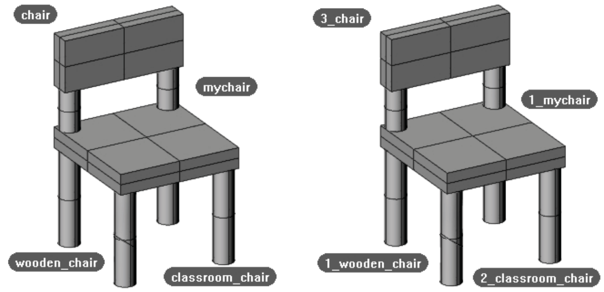


Fig. 3. Tagging a chair using a narrow and a broad folksonomy

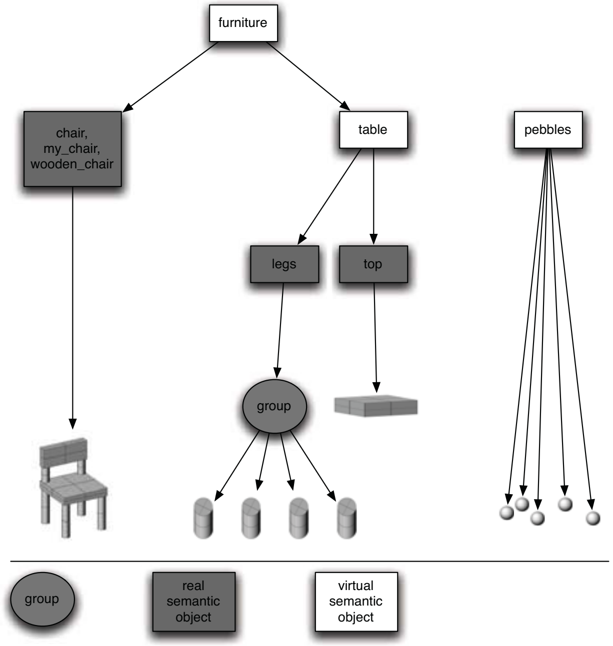


Fig. 4. Geometric and semantic objects

ing convergence towards a meaningful semantics. The issue will be further considered in the ongoing development of the project, where the users will experiment a prototype interface - under development - for tagging and their effort will be evaluated.

Concerning the accumulation of the tagging activity done by different users, the system may permit to store only one instance of a specific tag for a given object - as it happens in Flickr, the well-known web application for sharing photographs on the net - or also the number of occurrences. The structures derived from the latter approach are named *broad folksonomies*. They are opposed to the *narrow folksonomies*, that characterize the first approach, and - as explained in [19] - permit a better understanding of the terms that are more used by people for classifying objects.

Both approaches may be used for the 3D domain, as shown in Fig.3, where the chair on the left is tagged with a narrow folksonomy, while the same object - on the right - is tagged with a broad folksonomy. In both cases the system may preserve also the identity of the user tagging the object. Such additional information may enable additional processing, such as the extraction of tags assigned - for a given 3D world - by a single user or by a subset of users corresponding - for example - to a specific category.

Our design choice is to permit the accumulation of the instances for a given tag. Of course, a restricted folksonomy may be easily derived from the resulting broad folksonomy.

Another goal of our proposal is to maximize the number of existing 3D worlds on the net that may be enhanced with a semantic description. A parallel structure for storing semantic information is defined, that doesn't modify the existing relations stored in the scene graph. Such approach permits to enhance existing worlds, minimizing harms to the visualization and the interactivity that characterize the original 3D environments. The following section will show how the specification defined for tagging permits - taking advantage of the X3D standard - to reach such goal.

V. Associating tags to X3D worlds

We chose X3D as the target language for our methodology. X3D [5] is a widely diffused language for representing interactive geometric objects for the net. All the objects that are part of an X3D world - including the geometric objects - are described through *nodes* - that can be nested - and *fields* - where the properties of the objects may be stored. X3D represents the evolution of VRML97 and adds to it the capability to insert specific nodes for metadata, to specify information related to the objects of the 3D world.

Unfortunately, the X3D standard doesn't suggest how to take advantage of metadata nodes for defining structured semantic information inside 3D worlds.

In a previous work [11], the author suggested an approach for specifying high-level information for 3D worlds, using these nodes and an associated scene-independent domain ontology. In this work X3D metadata are used as the basis for associating tags to geometrical objects. This bottom-up approach is complementary to the previous one and is designed to be merged with it.

In the previous work we considered the concept of *geometric object* as opposed to the concepts of *real* and *virtual semantic objects*. The first category represents the raw information that may be found in any 3D file. It may be a single geometric shape or a group of geometric objects.

We coined the concept of real semantic object for all the cases where it is possible to associate an high-level meaning to a geometrical object.

Unfortunately, such association - as discussed in the previous sections - can't be always be found. There may be cases where geometric objects or groups defined in the scene graph can't be directly associated to a specific meaning, or such association doesn't make sense (e.g., many small objects - such as the stones displayed in Fig.4 - don't need a specific reference for each object, but they may collectively associated to a single label). Besides, there may be also the need to introduce higher-level semantic groupings for adding expressivity to the scene description.

For all those situations we defined the concept of virtual semantic object, a labelled container that collects a set of geometrical objects, lower-level semantic objects or even a mix of those entities.

In this work we take advantage of the same definitions. In spite of that, we propose a different complementary structure for metadata, suitable to the tagging needs, for giving the possibility to have different labels for the same object and for defining all the high-level semantics inside the same X3D file that describes the geometry.

Fig.4 show a sample of objects belonging to the three categories discussed above. Single geometric objects are evidenced through their geometric shapes. In some cases the geometric objects have been grouped by the world designer and this information - stored in the scene graph - has been evidenced with the circle labelled *group*. Real semantic objects are associated to geometric objects, single or grouped. Some of them are characterized by single tags (i.e., *legs* and *top*). The real semantic object associated to the shape that identifies the chair is characterized by a set of tags (*chair*, *my_chair* and *wooden_chair*), assigned by different users.

Virtual semantic objects, tagged as *table*, *furniture* and *pebbles*, have been specified where it has not been possible to use the existing shapes or grouping nodes of the scene graph for storing high-level information. These new objects are therefore introduced for completing the semantic description of the 3D world.

The code example displayed in Fig.5 shows the definition of a real semantic object, associated to the geometric shape defining the chair of Fig.4. A set of nested MetadataSet and MetadataString nodes are used for defining a metadata section inside the existing geometrical shape, *chair0123*.

All the tags and the number of occurrences for each tag are stored as a set of MetadataString nodes, nested inside a MetadataSet named *tagslist*. Another MetadataSet, named *grouping*, is used to contain the references to higher-level virtual objects; in this example the geometrical object is semantically associated - through the nested MetadataString node - to the virtual object named *furniture235*, tagged as *furniture* in Fig.4.

3D worlds are not made only of objects, but objects generate spaces that are inhabited by (virtual) humans. Such spaces may be proficiently labeled. That is the reason why - in coherence with what we did for the previous top-down proposal - in this work we extended the possibility to use tags also for spaces.

The X3D object that we currently use for associating tags to space is the ProximitySensor node, an invisible node that is used for monitoring the user action inside the 3D worlds. Proximity nodes may be used to define a set of locations and may also be nested for defining a hierarchy of spaces.

The code fragment displayed in Fig.6 shows how to associate tags to a proximity sensor available in the X3D scene. The structure of metadata nodes is similar to that one displayed

```
<Shape DEF='chair0123'>
  <MetadataSet>
    <MetadataSet name="folksonomy"
    reference="myfolksonomy">
      <MetadataSet name="tagslist" reference="">
         <MetadataString
         value="'0004' 'my_chair'"/>
         <MetadataString
         value="'0002' 'wooden_chair'"/>
         <MetadataString
         value="'0001' 'chair'"/>
       </MetadataSet>
       <MetadataSet name="grouping" reference="">
         <MetadataString name="furniture235"/>
       </MetadataSet>
     </MetadataSet>
  </MetadataSet>
 ...
 </Shape>
```

Fig. 5. A real semantic object tagged with three labels.

in the previous example. Also in this case different tags (i.e., *my_room*, *sitting_room* and *small_room*) have been used for classifying the same object. Because no higher-level space has been defined in the example, the MetadataSet node named *grouping* doesn't contain any MetadataString node.

```
<ProximitySensor DEF='room457'>
  <MetadataSet>
    <MetadataSet name="folksonomy"
    reference="myfolksonomy">
      <MetadataSet name="tagslist" reference="">
         <MetadataString
         value="'0004' 'my_room'"/>
        <MetadataString
        value="'0003' 'sitting_room'"/>
        <MetadataString
        value="'0001' 'small_room'"/>
      </MetadataSet>
      <MetadataSet name="grouping" reference="">
      </MetadataSet>
    </MetadataSet>
  </MetadataSet>
    ...
</ProximitySensor>
```

Fig. 6. A space tagged with three different labels.

The code given in Fig.7 illustrates how to define and tag a virtual semantic object starting from real geometric objects. The geometric objects are the components of the table presented in Fig.4. The virtual semantic object tagged as *table* is based on two different real semantic objects, defining the legs and the top of the table. Each real semantic object has a structure similar to that one described in Fig.5 and is linked to the virtual semantic object through the MetadataString nodes named *table457* (i.e., the identifier of the virtual object).

The virtual semantic object is defined as a set of MetadataSet and MetadataString, whose structure reflects that one adopted for real semantic objects. In spite of that, while the latter objects are defined inside existing geometrical and grouping nodes, the information related to virtual objects can't be referred to any existing node belonging to these categories.

For achieving our goal, we specify a section inside the

WorldInfo node, a standard X3D node used for giving a description of the content of a specific world. Each virtual semantic object - like the virtual object *table457*, tagged with the label *table* - is defined as a MetadataSet node, nested into the main MetadataSet named *virtual_objects*. The code shows also an additional relation of the virtual object *table457* with an higher-level virtual object, *furniture235*, not represented in the example.

```
<WorldInfo>
 <MetadataSet name="virtual_objects">
   <MetadataSet DEF='table457'>
    <MetadataSet name="folksonomy"
    reference="myfolksonomy">
      <MetadataSet name="tagslist" reference="">
         <MetadataString
         value="'0001' 'table'"/>
       </MetadataSet>
       <MetadataSet name="grouping" reference="">
         <MetadataString name="furniture235" />
       </MetadataSet>
     </MetadataSet>
  </MetadataSet>
 </MetadataSet>
</WorldInfo>
...
<Shape DEF='top0129'>
  <MetadataSet>
    <MetadataSet name="folksonomy"
    reference="myfolksonomy">
      <MetadataSet name="tagslist" reference="">
         <MetadataString
         value="'0001' 'top'"/>
       </MetadataSet>
       <MetadataSet name="grouping" reference="">
         <MetadataString name="table457"/>
       </MetadataSet>
     </MetadataSet>
  </MetadataSet>
 ...
</Shape>
<Group DEF='legs234'>
  <MetadataSet>
    <MetadataSet name="folksonomy"
    reference="myfolksonomy">
      <MetadataSet name="tagslist" reference="">
         <MetadataString
         value="'0001' 'legs'"/>
       </MetadataSet>
       <MetadataSet name="grouping" reference="">
         <MetadataString name="table457"/>
       </MetadataSet>
     </MetadataSet>
  </MetadataSet>
   ...
 </Group>
```

Fig. 7. Virtual and real semantic objects

VI. CONCLUSION

In this paper we have presented the results of an ongoing research activity targeted at tagging 3D worlds available on the net. The final goal of this research is to enhance low-level 3D information with semantic labels, for a full exploitation of 3D information available on the web.

In this work we focused on bottom-up folksonomic tagging, suggested as an approach complementary to the top-down ontology-based labelling, described in a previous work.

Currently the navigation and interaction potential of most 3D worlds is limited to what has been designed by the world author. Additional possibilities, such as advanced searching and filtering, may emerge from the availability of high-level information associated to the raw data. The use of a a widely diffused file format for the 3D web, X3D, and the specification of a unified methodology for tagging the components of the different worlds may extend these opportunities to a consistent number of 3D worlds deployed on the web, enabling cross-world searching, filtering and extraction of objects.

Ongoing work is focused on the implementation of a prototypical interface for verifying the design choices and receiving hints for future development.

REFERENCES

[1] "MPEG Homepage," http://www.chiariglione.org/mpeg/.
[2] F. Pereira and T. Ebrahimi, *The MPEG-4 Book*. Prentice-Hall, 2002.
[3] F. Nack and A. T. Lindsay, "Everything you wanted to know about mpeg-7 - part 1," *IEEE Multimedia*, vol. 6, no. 3, pp. 65–77, 1999.
[4] ——, "Everything you wanted to know about mpeg-7 - part 2," *IEEE Multimedia*, vol. 6, no. 4, pp. 64–73, 1999.
[5] X3D, "Extensible 3D (X3D) architecture and base components - edition 2 ISO/IEC IS 19775-1.2:2008," http://www.web3d.org/x3d/specifications/ISO-IEC-19775-1.2-X3D-AbstractSpecification/, 2008.
[6] I. M. Bilasco, J. Gensel, M. Villanova-Oliver, and H. Martin, "On indexing of 3D scenes using MPEG-7," in *Proceedings of the 13th Annual ACM International Conference on Multimedia*. ACM Press, 2005, pp. 471–474.
[7] P. Halabala, "Semantic metadata creation," in *Proceedings of CESCG 2003: 7th Central European Seminar on Computer Graphics*, 2003, pp. 15–25.
[8] H. Mansouri, "Using semantic descriptions for building and querying virtual environments," Ph.D. dissertation, Vrije Universiteit Brussel, 2005.
[9] RDF, "RDF Primer W3C Recommendation," http://www.w3.org/TR/rdf-primer/, 2004.
[10] OWL, "Web Ontology Language Guide," http://www.w3.org/TR/owl-guide/, 2004.
[11] F. Pittarello and A. De Faveri, "Semantic description of 3D environments: a proposal based on web standards," in *Proceedings of Web3D, 11th International Symposium on 3D Web*. ACM Press, New York, 2006.
[12] F. Pittarello, "Accessing information through multimodal 3d environments: towards universal access," *Universal Access in the Information Society*, vol. 2, no. 2, pp. 189–204, 2003.
[13] T. Vander Wal, "Folksonomy," http://vanderwal.net/folksonomy.html, 2004.
[14] B. Boiko, *Content management bible*. Wiley Publishing, 2004.
[15] L. Rosenfeld and P. Morville, *Information Architecture for the World Wide Web*. O'Reilly, 2006.
[16] P. Spyns, A. de Moor, J. Vandenbussche, and R. Meersman, "From folksologies to ontologies: How the twain meet," in *Proceedings of On the Move to Meaningful Internet Systems*, ser. Lecture Notes in Computer Science, vol. 4275. Springer, 2006, pp. 738–755.
[17] C. Van Damme, M. Hepp, and K. Siorpaes, "Folksontology: An integrated approach to turning folksonomies into ontologies," in *Proceedings of the ESWC Workshop Bridging the Gap between Semantic Web and Web 2.0*. ACM Press, New York, 2007, pp. 57–70.
[18] F. Carcillo and L. Rosati, "Tags for citizens: Integrating top-down and bottom-up classification in the turin municipality website," in *Proceedings of Online Communities and Social Computing: Second International Conference at HCI International*, ser. Lecture Notes in Computer Science, vol. 4564. Springer, 2007, pp. 256–264.
[19] T. Vander Wal, "Explaining and showing broad and narrow folksonomies," http://www.personalinfocloud.com/2005/02/explaining_and_.html, 2005.

TA-CAMP LIFE: INTEGRATING A WEB AND A SECOND LIFE BASED VIRTUAL EXHIBITION

Andrea De Lucia, Rita Francese, Ignazio Passero and Genoveffa Tortora

Dipatimento di Matematica e Informatica, Università degli Studi di Salerno,via Ponte don Melillo 1,Fisciano (SA),Italy
adelucia@unisa.it, francese@unisa.it, ipassero@unisa.it, tortora@unisa.it

Abstract:
Virtual Worlds are being ever more adopted for global commerce and in the future will be used in fields including retail, client services, B2B and advertising. The main advantage is the support provided to the user community in communicating while shopping.
This paper describes a project aiming at providing virtual exhibitions on Second Life: the TA-CAMP Life virtual expo system, which is the result of the integration between a web virtual expo and its extension on Second Life. The back-end web-based system supports the generation of an exhibition on Second Life and organizes the expo by distributing the exhibitors' stands on an island and enabling each stand to dynamically expose multimedia contents.

***Keywords**:* Second Life, virtual world, virtual expo, system integration, e-commerce.

1 INTRODUCTION

Enterprise marketing and external exchanges have new opportunities due to the development of network technologies. At the present, there is a growing interest towards 3D worlds which, thanks to the technological evolution, become more and more promising. Indeed, several worldwide organizations such as IBM and Linden are investing in this area.

In particular, Linden Lab proposes Second Life (SL) [8], the most popular Internet based Virtual World, if measured by number of subscribers and money exchange [4]. In Second Life, as well as in other Virtual World platforms, it is possible to interact with the other users, represented by avatars, through voice chat, text chat, and instant messaging. Virtual worlds offer a multi-tiered communication platform to collaborate and do business which provides the perception of awareness and presence that cannot be reached with e-mail, conference calls or other platform.

Linden Lab has created its digital currency for online exchange of goods and services: all processes of payment are virtualized and can be managed using the solutions offered by Second Life. At the present, several commercial organization are exploring SL world to support e-commerce. American Apparel, Adidas, Lacoste, Reedbok and Armani, for example, have opened virtual shops on Second Life. Indeed, 3D representations are suitable for electronic commerce because they emulate the shopping layout and many shopping items such as furniture, dresses, accessories, and so on. Exploiting the opportunity offered by the available virtual worlds enables organizations to obtain a simple setup that can be created with a reduced cost and can be accessed by a large number of people, without using specific input devices. In addition, during real life shopping customers often consult with each other about products. A multi-user environment offers to the users the possibility of collaborating while shopping, benefiting from each other experiences and opinions [9] [10].

In this paper we present the result of an ongoing project, the TA-CAMP project, which aims at providing the textile consortium of the Campania Region (Italy) with several services and, in particular, focusing on e-commerce and internationalization aspects. The TA-CAMP project offers a traditional web site to promote virtual expositions and also offers an enhanced version of this service on Second Life, named TA-CAMP Life.

In this direction we have proposed to the textile organizations of the Italian Campania Region a virtual expo enabling the textile consortium to reach a wide user community. In this way the marketing message is promoted in the traditional web world and also in an international economic context in continuous growing.

The virtual exhibition on SL is automatically generated creating a reception areas and several stands starting from the information available on the web version of the exhibition.

The rest of the paper is organized as follows: Section 2 introduces the main features of Second Life, while Section 3 describes how we have organized the exhibition on Second Life. Section 4 presents the architecture and the interaction modality offered by TA-CAMP life. Finally, Section 5 concludes.

2 THE SELECTED TECHNOLOGY

Several work have been proposed in literature aiming at supporting collaborative shopping, see [10] as an example. They often does not address scalability issues [9]. Peer-to-peer networked shopping CVEs have also been investigated in [9]. In this work peer-to-peer is preferred to virtual world based solution basing on the consideration that a group of buyer should contain at most dozen people in the same group at a given time. A virtual expo should host at the same time many avatars, distributed among the various stands and which can also be grouped all together during specific events, such as discussion meetings or awarding of prizes, scheduled by the expo organizations.

We decided to select Second Life to host the virtual expo features of the TA-CAMP project for several reasons, summarized as follows.

SL offers a tridimensional and persistent virtual world created by its "residents" in which is embedded a real economic system supporting the exchange of virtual goods and services.

SL hosts a community of over ten millions of users and concurrently, at each hour of the day, are on line about 100,000 users. Avatars can build and sell things, such as clothing or airplanes, and these transactions can be paid using the Linden dollar currency, which makes economic activities in the digital space directly connected to the earth-based economy.

It is important to point out that Linden Lab has estimated that the global market for virtual goods in a $1.5bn a year, and that Second Life residents bought and sold more than $360m worth of virtual goods and services in 2008 [6].

Users are represented by avatars and interact with the environment controlling the avatar actions. Second Life enables to use web, video, audio streaming, VOIP. People can privately as well as publicly chat on an open channel.

As investigated in [1], this movement in Second Life occurs in a natural manner and the user is able to control the events, he/she sees his/her avatar behaving as expected and the 3D world changing accordingly to his/her commands. Animations and gestures are offered to augment face to face communication. Once in the environment, people have a first person perspective, they participate, do not only watch. Situational awareness, "who is there", is well supported, as well as awareness on "what is going". Moreover, the user perception of awareness, presence and communication inducted by the environment are in general very positive [1].

SL offers the possibility to connect with external web-pages and internet resources.

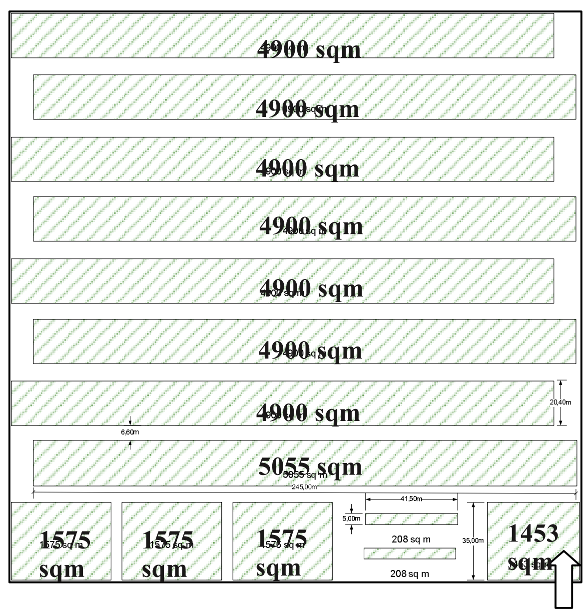


Figure 1. Space organization of the TA-CAMP Life island.

In addition, IBM and Linden Lab are developing together an in-house version of Second Life for businesses enabling enterprises to build secure virtual worlds that can be deployed behind a firewall [6]

Users access to SL using a client software that can be downloaded for free and is available for multiple platforms. Linden Lab, maintains a network cluster to host regions of 3D virtual environments, the “islands”. These islands contain user-created 3D content and can be interactively explored by the users that are logged into the system of SL. The content in SL is protected by a digital rights management system.

3. ORGANIZING A VIRTUAL EXHIBITION ON SECOND LIFE

SL is based on the archipelago metaphor, where space is organized in islands, which are connected each other via teleportation links, bridges, and roads. The island hosting the TA-CAMP life virtual expo has been designed as shown in Figure 1 and it is composed of a reception area and several stands.

The island has been designed in such a way to favor both the direct access to a stand of specific interest and to create a continuous path organized in pavilions. Pavilions are referred to a specific marketable goods and each one contains ten stands. Second Life enables to use web, video, audio streaming. People can privately as well as publicly communicate using Second Life chat or VOIP, collaborating while shopping.

3.1 The reception area

The exhibition offers a unique access area, signaled by an arrow in Figure 1. This area represents the reception of the exhibition where it is possible to consult the *Exhibitors’ Catalog*, structured in categories.

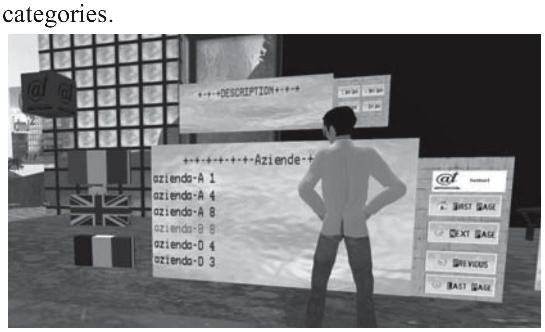


Figure 2.The Exhibitors’Catalog

In this environment several automatic distributors of expo’s gadgets, such as the shirt with the expo’s logo, are also available. A *Participant Detector Object* collects the data concerning the visitors and shows the number of participants.

The gadget distributor has been designed in such a way to collect the user preferences concerning models, colors and dress-material. The users answer to an implicit survey while customizing their gadget. A prototype of the *Survey/Gadget Distributor* is shown in Figure 3, where the user has chosen as a gadget a pair of trousers and is selecting their color and their material. The gadget is also decorated with the expo logo. Once terminated the gadget customization the user can wear or store it in his/her inventory.

The areas next to the reception are available to support exhibition events, such as presentations with a slide projector, as described in [3].

3.2 Stand Organization

Concerning the exposition stands, they are multi-part objects composed by at least 400 prims (the single building blocks of each Second Life object). Each island offers a limited number of prims, thus we have computed that each island can hosts at most 40 stands.

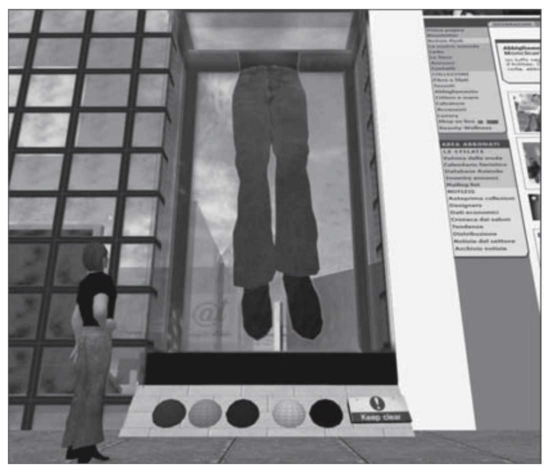


Figure 3.The Survey/Gadget Distributor prototype in the Reception Area

TA-CAMP Life offers two types of stand, depending on the availability of videos to be shown. In particular, it is possible to adopt a stand consisting of four areas (Presentation, Image, Video, Web) or a three areas one (Presentation, 2xImage, Web). In each area, in addition to the main communication channel (web, image or video), it is also possible to have an independent channel for audio diffusion. When several users are in an area and discuss using the chat, the text written by the user is saved in the web site back-hand for further analysis, provided the user permissions.

When a user accesses to a stand he/she goes in the Presentation Area, shown in Figure 4, where a web presentation of the firm and its products is displayed on a screen. He/she always can go back to the reception Hall using a teleport-like direct link.

In the Image Area it is possible to search the product catalog using the *Index Board* depicted in Figure 5 (a), to examine the detailed images of the selected product, Figure 5 (b), and its description displayed on a board adjacent to the image projector, as shown in Figure 5 (c). There is also the possibility of accessing to the front-end of the electronic commerce web site of the organization to order the item.

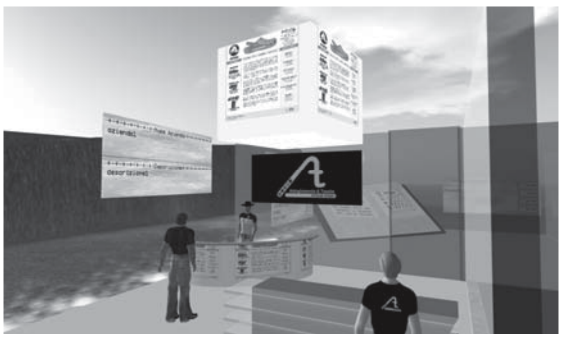


Figure 4.The stand Presentation Area

Another area is also available to provide video contents of advertisings or fashion-shows, for example. Also in this case an *Index Board*, depicted in the left-end part of Figure 6, enables to select the video to show by touching the related text line.

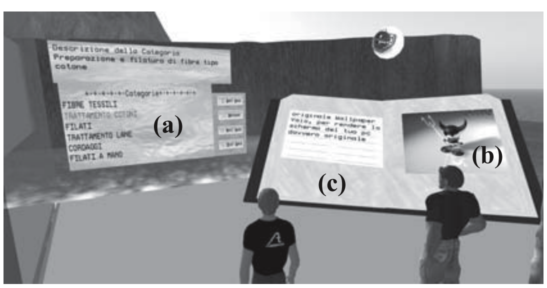


Figure 5. The Image Area

The web area is organized in a similar way and accesses to the expositor web site.

4 INTEGRATING AN IN-WORLD AND OUT WORLD VIRTUAL EXHIBITION

The project requirements established that the virtual exhibition had to be offered in both web and Second Life modalities and that, once populated the database of the web version, the SL exhibition had to be automatically generated.

Users who access to the web version of the expo using the browser are also invited on the expo in Second Life. By clicking on a Second Life link they are teleported in the expo areas of TA-CAMP Life island.

Figure 6. The Video Area

The identified actors are:

- *The system administrator*, managing the virtual exhibition. This includes the definition of the start and finish date of the exhibition, the association of an exhibitor to a stand, the exhibitor access right definition, etc. These functionalities are supported by a web interface.
- *The exhibitor*, managing his/her stand and the contents to be shown.
- *The customer*, visiting the expo and buying goods through the expositor e-commerce web site.

The web and virtual world interaction modalities collect the needed contents inquiring a common *Content Management System (CMS)*, as illustrated in Figure 7, where an overview of the TA-CAMP Life system architecture, with the different components distributed over several servers, is shown.

The *SL Expo objects*, such as projectors and Index boards, are resident on the *Second Life Linden External Server*. All these objects expose an active behavior obtained by using the programming language offered by SL, namely Linden Scripting Language [5].

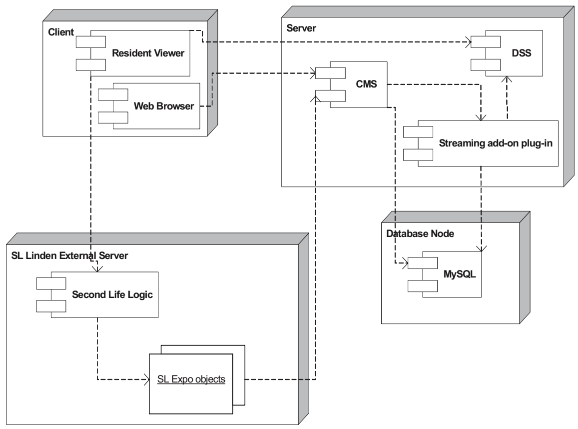


Figure 7. The system architecture

The virtual exhibition is dynamically generated collecting the data offered by the traditional web site. In particular, each *SL Expo object* is dynamically populated as follows: the *SL Expo object* sends a *HTTP* request to the *CMS*, specifying its stand identifier, to obtain the appropriate content to be displayed in Second Life. The *CMS* embeds the required information in a *HTTP* response towards the considered *SL Expo object*. This mechanism enables to get a new 3D exhibition each time the web site starts a new expo. An ad-hoc developed plug-in of the *CMS*, named *Streaming add-on plug-in*, communicates with a *Darwin Streaming Server (DSS)* component, integrated into the system to provide streaming capability to both the CMS and Second Life. It also provides the possibility to access, in a controlled manner, to a variety of multimedia contents from an exhibition stand.

4.1 Accessing to multimedia contents from Second Life

In this sub-section we describe how TA-CAMP Life accesses, in a controlled manner, to a variety of multimedia contents during a visit of the Video Area of a stand.

It is important to point out that SL technology exposes to land owners the availability to connect each land parcel to media content which can consist of images, videos, audios or web pages. To exploit this feature, the multimedia materials have to be stored on an external server.

Communication involving the objects and the external world has been performed using *HTTP* requests/responses, while intra-object communication relies on link or chat messages. A *link message* is adopted when sender and receiver are embodied in the same composite object. *Chat messages* may be exchanged among several objects in the same island. Different kinds of chat messages can be selected, depending on the sender and receiver distance. In addition, each chat message can be sent on a reserved channel in such a way to have a unique receiver [7].

To enable customers to play video contents during their visit, the Video Area is equipped with the *Content Index Board*, which displays the catalog of the multimedia contents associated to a specific stand. Once selected a content, it is played on the *Content Board.* Figure 8 shows the *In and Out World Objects* involved to display multimedia contents on the *Content Board* in Second Life. The *Content Index Board* exposes two objects: the *Page Button* to go forward and backward in the content list and the *Content Selector* object. The *Content Index Board* requires the content details to the *Stand object* which enquires the *Content objects*. The *Stand object* returns this information to the *Content Index board* for displaying it. The *Content Selector* highlights the index element selected by a *touch action* and communicates its position to the *Content Index Board*, which, in turns activates the *Content Board*. The latter sends a *HTTP* request to the identified resources out world.

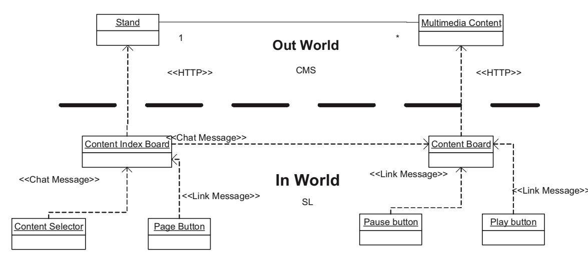


Figure 8. In/out world communication in the Video Area

SL enables to show text only in terms of images. Chats can also be used to display textual information, but they are not suitable to show large text. Thus, to display textual contents on the boards we adopted a library, the *XyzzyText* library [13], enabling to create special elementary prims able to display a pair of letters on each face. By disposing these elementary blocks on the surface of a board it is possible to show multi-line text. As an example, to show the content list on the *Content List Board* the text to be displayed is required to the *Stand object* out world and then arranged using the *XyzzyText library*. An example of board is shown in Figure 2, where the exhibitors' catalog board capable of displaying a matrix of 10 x 40 characters is depicted.

5 CONCLUSIONS

In this paper we have described the main features of the virtual exhibition components of the TA-CAMP Life project, enabling two variants of a virtual expo, one web-based and the other based on the Second Life virtual world, to coexist. Using a unique database for both the approaches a complete virtual world expo can be automatically generated. It is important to point out that TA-CAMP Life does not replicate the remoteness loneliness of an exhibition. Even if it offers a product catalog as in its web version, TA-CAMP Life also promotes the social texture of a real exhibition, along with the collaborative nature of buying and offers to the exhibitors the possibility of organizing synchronous events. The system also provides survey features and sensors to examine the user behavior and collect information useful to foresee marketing trends. We plan to use this information together with other data concerning the user behavior for anticipating the his/her needs in forthcoming interactions, investigating the differences between the adaptation in a multi-user environment and similar approach proposed for single-user environments, such as [1][2]. Future work will also be devoted to investigate how to adopt the functionalities offered by SL for controlling the avatars, integrating virtual agents into TA-CAMP Life. In this way it will be possible to support customer care, following the directions traced in [10].

ACKNOWLEDGEMENTS

This research has been supported by Regione Campania, founding the TA-CAMP project.

REFERENCES

[1]. Bonis, B., Stamos, J., Vosinakis, S., Andreou, J., Panayiotopoulos, T., (2007), Personalization of Content in Virtual Exhibitions, in the Proceedings of the 2nd international conference on Semantics And digital Media Technologies:pp. 172-184

[2]. Celentano, A., Pittarello, F., (2004), Observing and Adapting User Behavior in Navigational 3D Interfaces, in the proceedings of the working conference on Advanced visual Interfaces, Gallipoli, Italy, pp. 275 – 282.

[3]. De Lucia, A., Francese, R., Passero, I., Tortora, G. (2009) Development and Evaluation of a Virtual Campus on Second Life: the case of SecondDMI. Computer & Education. Elsevier. Vol. 52, Issue 1, January 2009, Pp. 220-233, doi:10.1016/j.compedu.2008.08.001.

[4]. Edwards, C., Another World. IEEE Engineering & Technology, December 2006.

[5]. Linden Scripting Language, http://wiki.secondlife.com/wiki/LSL_Portal.

[6]. Nichols, S., IBM to build corporate Second Life Virtual worlds for enterprises, http://www.vnunet.com/vnunet/news/2213531/ibm-build-private-second-life.

[7]. Rymaszewski, M, Au, W., J., Wallace, M., Winters, C., Ondrejka, C., Batstone-Cunningham, B., Rosedale, P., Second Life: the office guide. Wiley Press, 2007.

[8]. Second Life. http://secondlife.com.

[9]. Khoury, M., Shirmohammadi, S., Accessibility and Scalability in Collaborative eCommerce Environments.I nternational Journal of Product Lifecycle Management. Vol. 3, pp. 178 – 190, 2008.

[10]. Shen, X., Shirmohammadi, S., Desmarais, C., Georganas, N.,D., Kerr, I., Enhancing e-Commerce with Intelligent Agents in Collaborative e-Communities, In the Proceedings of the IEEE Conference on Enterprise Computing, ECommerce and EServices, San Francisco, CA, U.S.A, IEEE, June 2006.

[11]. Shen, X., Radakrishnan, T., Georganas, N., vCOM: Electronic commerce in a collaborative virtual world. In Electronic Commerce Research and Applications 1 (2002), ELSEVIER, 2002, pp. 281-300.

[12]. Williams, I., Linden Lab expands e-commerce in Second Life, http://www.vnunet.com/vnunet/news/2234753/linden-lab-expands-ecommerce.

[13]. XyzzyText, http://wiki.secondlife.com/wiki/XyzzyText

Genòmena: a Knowledge-Based System for the Valorization of Intangible Cultural Heritage

Paolo Buono, Pierpaolo Di Bitonto, Francesco Di Tria, Vito Leonardo Plantamura

Department of Computer Science – University of Bari
Via Orabona 4, 70125 Bari
{buono, dibitonto, francescoditria, plantamura}@di.uniba.it

The Italian nation is famous for its history and cultural heritage. Artefacts and cultural treasures dating back to various periods of the past are often preserved in museums, but traditions, dialects, cultural events are some examples of intangible heritage from the past that cannot be kept in museums. They are the basis of current cultures but nevertheless the historical memory of them tends to disappear since it is difficult to preserve for the new generations. In this paper we present Genòmena, a system that has been designed to store and preserve intangible cultural heritage, thus saving it for posterity. Genòmena allows different types of people to access such intangible heritage via a Web portal. Thanks to its underlying knowledge-base, it is possible to gain information in different ways, like multimedia documents, learning objects, event brochures.

1. Introduction

The variety of people's cultures is the result of a long evolution that, during the course of centuries, transforms a territory and the customs and traditions of its inhabitants. History is not only written in great literary works, but is also preserved through traditions, dialects, etc., which all contribute to people's culture and cultural heritage. Only through the study and preservation of this heritage can the memory of a territory and its inhabitants be kept alive and appreciated in the present time.

The 2003 Convention for the Safeguarding of the Intangible Cultural Heritage defines the intangible cultural heritage as "the mainspring of our cultural diversity and its maintenance a guarantee for continuing creativity" [1]. Intangible cultural heritage is manifested in domains such as: oral traditions and expressions, including languages and dialects as a vehicle of the intangible cultural heritage; performing arts, e.g. traditional music, dance and theatre; social practices, rituals and festive events; traditional craftsmanship.

Since the time of "Magna Grecia" (8th century BC), Italy, and especially the Puglia region, has been a crossroad of peoples coming from the Mediterranean basin (and not only). Puglia underwent several periods of foreign domination and was the site of many important pilgrimages to visit the relics of Saint Nicholas, one of the most revered saints of all Christendom. Museums and archeological parks preserve much of the ancient heritage, but traditions, dialects, cultural and religious events are examples of intangible heritage that it is difficult to maintain for future posterity [2].

The Genòmena system has been developed to preserve and recover the ancient traditions of the people of the Puglia region. The name derives from the ancient Greek word γενòμενα, which means events. As will be described in the paper, the system provides information to various types of people and in different ways, namely multimedia documents, learning objects, event brochures.

Genòmena has three main objectives: 1) to foster the dissemination of intangible heritage in order to keep its historical memory alive; 2) to promote tourism in the Puglia region by providing detailed information about items of intangible heritage; 3) to support research on cultural heritage.

One of the peculiar features of Genòmena is that it offers the possibility of performing very advanced data searches. In fact, the underlying knowledge base makes it possible to retrieve information on the basis of semantic as well as spatial and temporal relationships among the stored objects.

The paper has the following organization. Section 2 briefly describes related work. Section 3 presents the system architecture and the main users of Genòmena. Section 4 describes our novel approach to help users to find relevant information, based on an ontological representation and on a knowledge-based search agent. Finally, some conclusions are reported.

2. Related work

Genòmena is a novel system that, among its various goals, aims at supporting the exploration of relationships among several cultural heritage documents. Other systems have been built for this purpose. PIV is a system that allows users to search for documents related to Pyrenean cultural heritage [13]. PIV is based on Web services and allows people to retrieve documents according to a geographic search. It is equipped with both a content-based search engine and a semantic engine. The semantic engine is integrated with a geographical database that is able to search for spatially related documents. The results are visualized using a cartographic representation in which each document is represented by a point near the place it

evokes. The system do not retrieve temporally related documents (e.g. documents written in the same period).
The P.I.C.A. project aims at preserving and valorizing the Po Valley and the Western Alps [14]. The system has been developed in order to allow users to access cultural documents related to this territory. It is equipped with an XML-based search engine that retrieves documents by using both traditional keywords based searches and graphic maps. The extracted documents are visualized as cards describing specific items (e.g. monuments). Graphic maps show topographic information, thanks to interaction with the MapServer. Also in this case, the user can only browse documents according to spatial criteria.
An interesting system is T.Arc.H.N.A., that provides cultural contents by a narrative visualization of items [20]. The narrations, composed by XML files and visualized as multimedia contents, are searched for by Archeologist, using a Narration Builder, which is a search engine that generates queries to be sent to different databases, containing documents about Etruscan cultural heritage.
Meyer et al. introduce the Virtual Research Environment, a Web-based search engine that allows users to perform spatial and/or temporal explorative analyses [15]. This engine is able to perform advanced searches, creating queries that combine temporal and geographic criteria. This system allows users to perform studies of the history of a territory and a virtual visit of a site. Lastly, the search engine provides keywords and images based searches, since all the multimedia objects are described by metadata. The visualization of the retrieved documents is based on both interactive maps, which allow a virtual exploration of a territory, and 3D models, that allow access to documents referencing a given place at a given period of time. However, the data are stored in relational databases and XML files, and there is no ontological representation of the domain of interest, preventing semantic searches.
The semantic search is based on explicit knowledge representation and can reveal every kind of relationship by using inferential processing.
Knowledge-based search engines use their knowledge about the user and items in order to generate suggestions, by reasoning on which items satisfy the user requests. These systems fall into two categories: rule-based, and case-based.

Figure 1 The visualization of search results in Genòmena

The rule-based engines use a set of rules to infer correlations among different items. The case-based engines are a subset of the former and adopt the assumption that a new problem can be resolved, by retrieving and fitting the solution found for already stored similar cases.
An example of a rule-based engine can be found in [18], where an expert system gives search results about hotels, providing the reasons for the selected items. Instead, an example of a case-based engine can be found in [19]. In this paper, the authors describe the Entree system, which is able to suggest restaurants. On the basis of the information inserted by the user, the system selects from its knowledge base a set of restaurants that satisfy the user preferences. Finally, the system sorts the retrieved restaurants according to their similarity with the current case. The Genòmena system acts as a rule-based engine.
The visualization of data that have inherent spatio-temporal information in the Web is not an easy task. This is confirmed in the study performed by Sutcliffe et al. [16]. Several ways of presenting results of a query have been adopted. Yee et al. propose a visualization based on facets [17]. We were inspired by this work and the presented results of the advanced search engine using a multidimensional approach. The visualization is dynamic and provides the possibility to apply filters. Figure 1 shows a dynamic web page, split in three areas. The top left area contains information on the search engine subdivided into general items, geographic areas, time, respectively. The bottom left area represents a tree that contains IICH documents, brochures, learning objects that are correlated according to the search results presented in Section 4. The right area presents the details of the retrieved (and filtered) items.

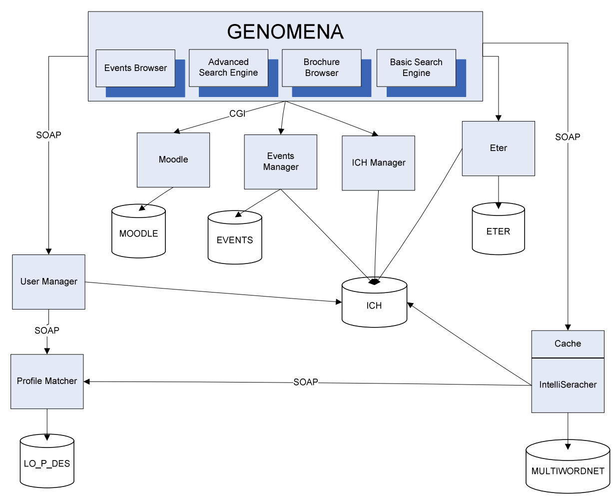


Figure 2 Genòmena system

3. System architecture

Genòmena is a modular, distributed system that includes web applications, web services and several databases.
As shown in Figure 2 the main entrance of the system is the Genòmena portal, which allows people to access the Events browser, the search engine, and the Brochure browser.
Information on the system is shown according to different user permissions, managed by the User Manager Web service. The system provides an advanced search only to

registered users, and allows content management only by system administrators or cataloguers, as will be seen later in the paper.

The Advanced Search Engine finds relationships among Items of Intangible Cultural Heritage (IICHs). In order to produce the search results it interacts with IntelliSearcher, which is a knowledge-based search engine whose aim is to find items that are related by semantic relationships. Personalized search results are provided by IntelliSearcher, exploiting the Profile Matcher, which assigns a score to each resource found, according to the user profile.

Registered users may get information not only as IICH documents and event brochures but, since the system also manages learning objects on topics related to intangible cultural heritage, they can also access on-line courses provided by the Moodle web application. Teachers organize these courses by assembling a set of learning objects, imported by eTER as a web application that permits the upload of learning objects described by metadata, based on IEEE LOM [5] and fEXM [11].

The cataloguer inputs all IICH data through the IICH manager web application.

Events Manager is a decision support system that assists event organizers in planning events. Events data are stored in the IICH database through an ETL process [12].

3.1. Genòmena users

Genòmena is a system designed to manage items of intangible cultural heritage, in order to preserve their memory. Thus, its main objective is to support the dissemination of stored information to all citizens, ranging from school children to senior people. Genòmena is also a great source of information for researchers working on cultural heritage and is intended to support tourists visiting the Puglia region.

The users accessing the system are very different, and interested in getting information on different aspects of the same item. For example, a student interested in the traditions of his own territory can access learning objects, which explain information about a certain item by using a didactic approach; the tourist, who is interested in cultural aspects related to religion, gastronomy, etc., gets information about events such as trade fairs, religious events, shows, and can access brochures concerning the requested event; the researcher, who might be interested in getting anthropological and/or philological data, can review documents, and technical material, written according to the Italian Central Institute for Cataloguing and Documentation (ICCD) standard, which contains useful details [3, 4].

In order to adequately support users' requests, all the available material must be stored and organized in a structured way, in order to facilitate their retrieval and fruition.

The main users of Genòmena, who work with the system for either inputting data or for retrieving them, are the following. The *cataloguer*, who is very familiar with the ICCD standard and inputs data describing an IICH according to this standard. The *researcher*, who is interested in items related to history and cultural heritage. As we have said, the main objective of Genòmena is to disseminate knowledge about intangible items. Therefore, other users of the system are the *local inhabitants*, who are interested in information about local events, religious traditions, multimedia items like photos, video and oral stories. The catalogued items can also be an object of study by *school children*, who are mainly interested in short on-line courses related to history, religion and their connections with the territory. For this purpose, the system is integrated with a Learning Management System (LMS) which manages learning resources, related to the most important cultural items classified in the repository. The e-learning environment increases the possibility of sharing of the resources, providing on-line courses to be accessed at any time from any location. Such courses are organized by *teachers* working with the Open Source e-learning platform *Moodle*, which is integrated in the Genòmena system.

Tourism promotion is strictly related to cultural dissemination. Tourists may benefit from cultural items and plan customized paths in order to improve their knowledge about the habits and the traditions of the cities they want to visit.

There are different kinds of *tourists*. The business traveler typically looks through images, searching for event schedules, city maps and traditional cooking. Other popular tourists in Puglia are those interested in religion, since the region is full of important churches and religious monuments. Such tourists are mainly interested in paths and journeys proposed by church organizations.

Genòmena also provides the possibility of organizing special events related to IICHs through the Event Manager module, used by *event organizers*. Finally, there are other people that work behind the scenes. Specifically the *system users* that maintain the whole system.

Genòmena has been designed to support all these user categories, which have been analyzed in depth in order to develop a system that supports their needs and expectations, according to a user-centred approach.

Users can search for content and browse several types of documents. Currently, three types of documents are supported: multimedia documents structured according to the ICCD standard for describing an IICH (called *IICH document* in the rest of the paper), event brochures, and learning objects. These documents can be accessed in different ways, each providing contents with different media.

4. Finding relevant information

In order to represent the information about intangible cultural heritage in the system, an in depth study of the domain was conducted in collaboration with cultural heritage experts.

As shown in Figure 3, the system knowledge base distinguishes three types of knowledge: factual, specific, and general.

The factual knowledge describes different items of cultural heritage and is stored in the database of the system. Examples of the factual knowledge are IICH documents, Learning Objects, event brochures.

The specific knowledge describes the geographic and historical context of the single item of factual knowledge, providing specific spatial-temporal relationships. It is represented in ontological form according to OWL syntax. For instance the IICH document about the relics of Saint Nicholas is related to the history of the Saint.

The general knowledge is the basic knowledge used to build specific knowledge in order to carry out the inference process within the KB. The general knowledge describes the historical context of the specific knowledge and is represented in ontological form. For instance, the specific knowledge about the saint's history is contextualized in the history of Christianity, or the specific knowledge about different people's traditions is contextualized in the history of the people. The general knowledge, in the example, covers a period that goes from the Christian period to the present day and represents traditions, cultures, dominations, religions. The representation knowledge used by the system for providing suggestions to users is reported in detail in the next section. The system knowledge is formalized in order to explain how it can provide suggestions for searches.

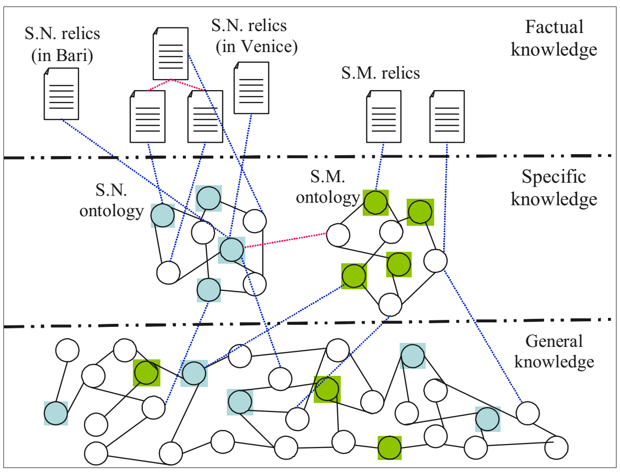


Figure 3 Three kinds of knowledge involved in the intelligent search process

4.1. System knowledge

The factual knowledge consists of IICH documents, learning objects and event brochures. The factual knowledge objects are shown in Figure 3. Each intangible cultural heritage document contains data structured according to the ICCD standard and is stored with learning objects and event brochures. An IICH document describes an item of intangible cultural heritage and is composed of the following macro-descriptors: (a) *codes*, that represent the identifiers of the items at regional level; (b) *definition*, that contains the description of the item and its membership category; (c*) geographic location*, that describes where the item is located, specifying nation, region, province, and city; (d) *time period*, that indicates the period of the year when the item happens; (e) *relationships*, that contain the references to the related items; (f) *analytical data*, that contain a detailed description of the item; (g) *communication*, that describes the kind of communication (such as vocal and/or instrumental) that accompanies the item; (h) *individual actor*, that indicates the presence of a single person in the item (for instance a ballad singer that tells the traditional tales); (i) *joint actor*, that indicates the presence of a set of people with their respective roles (for instance in a historical procession there are a lot of actors, such as knights, jugglers, tumblers, and so on); (j) audio/video/photo document, that stores the links and the descriptions of the multimedia content related to the item; (k) *element specification*, that contains further information about the item; (l) *data access*, that points out the item copyrights; (m) *writing mode*, that stores the name of the expert cataloguer of the item and the date of cataloguing; (n) *features* that are indicative of the kind of events related to the item.

An example of IICH is "the over the sea procession of Saint Nicholas' statue". In the system, this IICH is represented according to the ICCD standard. In this case, only eight of the sixteen descriptors are necessary. Specifically, this item has the following descriptors activated. Code (a): 1601000005. Definition (b): "vessel's statue procession in the sea". Geographic location (c): there are various details such as country, city, etc. In this case country is Italy and city is Bari. Time period (d): May 7th. Analytical data (f): in this section there is a long description about the intangible cultural heritage item. Element specification (k): Rituals and traditional festive events. Access data (l): no privacy or security limitation. Writing mode (m): Archive.

As regards the learning objects, they are described using IEEE LOM [5]. The event brochures are described by the name of the event, the schedule of sub-events, sponsors supporting the event organization and the mass media advertising the event. Each learning object and event brochure refers to one or more IICH documents so in the search process the system finds not only an item of IICH document but also related learning objects and brochures.

The system represents the specific and general knowledge using the same representation model, based on objects, with properties and relationships, using OWL language [6]. In particular, the relationships are expressed in terms of time and space.

The spatio-temporal representation has raised several research questions, for instance: how to define the same religious worship that takes place in different times and in different geographic areas; how to define the same title borne by different persons, i.e. the king of France is represented by different people, according to the specific moment in time we are considering, and so on. The problem has been solved by using the event calculus, an evolution of the situation calculus, which permits an event to be considered as a spatio-temporal portion [7]. Using this technique it is possible to generalize the concept of event as a space-time portion rather than just en event in time. A set of functions, predicates and rules was thus defined, on which space-time reasoning is based. For instance the following definitions have been made:

- Occurrence(e, t): that indicates that the event e occurred at time t
- In(e1, e2): that indicates the spatial projection of event e inside another space (e.g. In(Rome, Italy)
- Location(e): that indicates the smallest place that completely covers event e (e.g. Location(relicX) = ChurchY
- Start(): that indicates the first moment of time of the event

- End(): that indicates the end of the event
- Consecutive(i,j)⇔((Time(End(i))=(Time(start(j))): that establishes that two events are consecutive if the instant when the first one ends is the one when the second one starts.

These predicates and functions allowed us to define the relations highlighting analogies among fragment of knowledge. There are three different types: time, space and concept.

4.2. The search process

The knowledge representation is used by the system in order to suggest relevant contents that are related to the user's query, which is a string inserted by the user.
The search process is composed of three main phases:

1. Lexical enrichment of the search string: the string inserted by the user is parsed and completed using the lexical database MultiWordNet [8, 9]. In this phase the query string is tokenized and formatted for the information retrieval process. The terms in the query string are enriched with synonyms taken from the MultiWordNet database.
2. Search and selection of the relevant IICH documents: starting from the enriched query string, retrieved from the factual knowledge. For each term of the string, a list of IICH documents, ranked by relevance, is produced.
3. Suggestions: the system computes correlations of each selected IICH document with other IICH documents, using the specific and general knowledge.

In the suggestion phase, thanks to the information on IICH documents found, together with the specific and general knowledge represented via event calculus and the ontology (stored in OWL format), a run-time knowledge based is generated. Concepts, instances and properties of the ontology needed to be formalized in declarative language: in particular, a hierarchical representation of the concepts and the properties of the ontology is stated as rules. The instances are inserted in the database in the form of facts. After creating the database, the goals for determining the IICH documents to be suggested were defined. In this way it is possible to combine various types of relations (e.g. contemporary, neighboring events, …) in order to suggest the most relevant IICH.
The result of this process is a list of IICH documents, which have spatial and temporal relationships according to the initial search string. Moreover, using the relationships in the factual knowledge, the system provides a list of learning objects and event brochures related to the retrieved IICH documents. The output is then organized by the profiling system, that ranks and orders the results according to the needs of the specific user interacting with the system.

4.3 Inferring process: an example

In order to understand how the relationships among the objects are used in the inferring process, an example of the knowledge base is presented, according to the knowledge representation shown in Figure 3.
An IICH document about Saint Nicholas relics is contextualised in the ontology that describes the life and the work of the Saint. The specific knowledge is contextualised in the time-space dimension in the general knowledge.
Let us suppose that a cultural heritage researcher, interested in the history of Saint Nicholas, defines as a search criterion the following string: "relics of Saint Nicholas".
The system initially finds the item of IICH document related to the search string in the factual knowledge. On the basis of the data contained in the retrieved IICH document, the following facts are asserted in the knowledge base and added to the ontology (of Saint Nicholas) in the specific knowledge:

1. In 1087 sailors of Bari stole some of the bones of Saint Nicholas
2. In 1100 sailors of Venice stole other bones of Saint Nicholas
3. Some bones of Saint Nicholas are in San Niccolò Lido Church
4. Some bones of Saint Nicholas are in Saint Nicholas Cathedral

Moreover, in order to join the specific knowledge about Saint Nicholas with the other specific knowledge the inferring process uses the general knowledge. In the example, the following facts are asserted:

1. San Niccolò Lido Church is in Venice Lido
2. Saint Nicholas Cathedral is in Bari
3. Saint Nicholas is patron of Bari
4. Venice Lido is in Venetian territory
5. San Marco is patron of Venice
6. San Marco Cathedral is in Venice
7. San Marco relics are in San Marco's Cathedral

Thanks to this process, the system can make the following logical deduction:

San Marco and Saint Nicholas are correlated

These new inferred facts represent the result of the inferring process. In this way, the system shows the IICH document related to both San Marco relics and Saint Nicholas relics because both of them are kept in churches that are spatially close.??

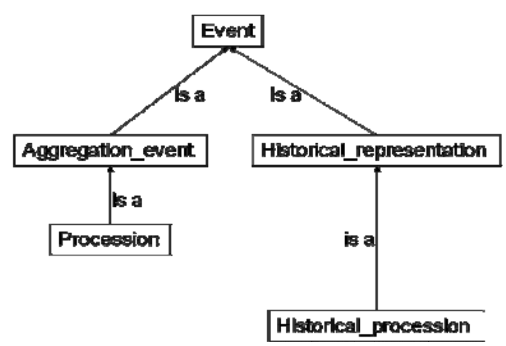


Figure 4 Part of ontology describing a religious event

For a better understanding of the working logic let us suppose that a user finds an IICH document referencing an event related to the life of a Saint and that the user is interested in further events that happen in the same moment as this event. Two kinds of temporal relationship have to be considered: the first defines the exact matching of two or more events during time; the second defines the temporal analogy among past events. For example, on the 6th of December of every year, there is the celebration of Saint Nicholas. On the basis of the first temporal relationship, the user finds further cultural events that happen in the same period of the year. On the other hand, thanks to the second relationship, (s)he is also able to find events like the old winter celebration, that, some centuries ago, happened exactly on the 6th of December [10]. The added value of the knowledge based search consists of semantic relationships discovered automatically. In Figure 4, the class diagram shown reports a part of the ontology.

Conclusions

This paper has presented the Genòmena system, which is designed to manage intangible cultural heritage and to support its preservation and valorization in order to keep alive the memory of a territory and its inhabitants. Indeed, one of the main novelties of Genòmena is its search engine, that exploits ontological representations and makes it possible to perform advanced searches, so that information is retrieved on the basis of various relationships among the stored objects. Moreover, the system uses a semantic engine that is able to find spatial, temporal and categorical relationships among items of intangible cultural heritage. The results are presented using a multidimensional dynamic Web interface that allows users to refine the output and analyze a subset of retrieved documents.

Acknowledgement

This work is supported by the Genòmena grant, provided by the Puglia Region. We would like to thank Prof. Maria F. Costabile and Prof. Teresa Roselli for the useful discussions during the development of this work. We also thank the students N. Policoro, M. Gadaleta, G. Vatinno, and M. T. Facchini for their contribution to the system implementation.

References

[1] http://www.unesco.org/culture/ich/ UNESCO Web site about Intangible Cultural Heritage. Last access on March 2009.

[2] Lupo E., Intangible cultural heritage valorization: a new field for design research and practice. International Association of Societies of Design Research, Emerging Trends in Design Research. Hong Kong Polytechnic University, 12-15 November 2007.

[3] Aiello A., Mango Furnari M., Proto F., ReMuNaICCD: A formal ontology for the Italian Central Institute for Cataloguing and Documentation, Applied Ontology, vol. 3, 2006.

[4] www.iccd.beniculturali.it, Central Institute for Cataloguing and Documentation, last access on March 2009.

[5] Learning Technology Standards Committee of the IEEE. Draft Standard for Learning Object Metadata in IEEE-SA Standard 1484.12.1, http://ltsc.ieee.org/wg12/ files/LOM_1484_12_1_v1_Final_Draft.pdf. Last access on March 2009.

[6] McGuinness D. L., van Harmelen F., OWL Web Ontology Language Editors, http://www.w3.org/TR/2004 /REC-owl-features-20040210. 2004. Last access on March 2009.

[7] Russell S. J., Norvig P., Artificial Intelligence: A Modern Approach. Prentice Hall. NJ: Upper Saddle River. 2003.

[8] Pianta E., Bentivogli L., Girardi C. (2002). MultiWordNet: Developing an Aligned Multilingual Database. Proc. of the First International Conference on Global WordNet. Mysore. India, 21-25 January, pp. 293-302.

[9] Bentivogli L., Forner P., Magnini B., Pianta E. (2004). Revising WordNet Domains Hierarchy: Semantics, Coverage, and Balancing, Proc. of COLING 2004 - Workshop on Multilingual Linguistic Resources. Geneva. Switzerland, 28 August 2004, pp. 101-108.

[10] Jones C. W. (1978). Saint Nicholas of Myra, Bari, and Manhattan: Biography of a Legend. Chicago and London: University of Chicago Press.

[11] Roselli T., Rossano V. (2006). Describing learning scenarios to share teaching experiences. International Conference on Information Technology Based Higher Education and Training. IEEE Computer Society Press. Sydney. Australia. 10-13 July 2006, pp. 180-186.

[12] Kimball R. (2004). The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data. John Wiley & Sons.

[13] Marquesuzaà, C., Etcheverry., P. (2007). Implementing a Visualization System suited to Localized Documents. Fifth International Conference on Research, Innovation and Vision for the Future. P. Bellot, V. Duong, M. Bui, B. Ho (eds.). SUGER, Hanoi. Vietnam. 05-09 March 2007, pp. 13-18.

[14] Agosto E., Demarchi D., Di Gangi G., Ponza G. (2005). An open source system for P.I.C.A. a project for diffusion and valorization of cultural heritage. CIPA 2005. XX International Symposium On International Cooperation to Save the World´s Cultural Heritage. Torino, Italy. 26 Sept. - 1 Oct. 2005, pp. 607-611.

[15] Meyer E., Grussenmeyer P., Perrin J. P., Durand A., Drap P. (2007). A web information system for the management and the dissemination of Cultural Heritage data, *Journal of Cultural Heritage*, vol. 8, no. 4, Sept. - Dec. 2007, pp. 396-411.

[16] Sutcliffe, A. G., Ennis, M., and Watkinson, S. J. (2000). Empirical studies of end-user information searching. *Journal of the American Society for Information Science.* Vol. 51, no.13, (Nov. 2000), 1211-1231.

[17] Yee, K., Swearingen, K., Li, K., and Hearst, M. (2003). Faceted metadata for image search and browsing. Proc. of the SIGCHI Conference on Human Factors in Computing Systems CHI '03. Ft. Lauderdale, Florida, USA, April 05 - 10, 2003. ACM, New York, NY, 401-408.

[18] Gobin B. A., Subramanian R. K. (2007). Knowledge Modelling for a Hotel Recommendation System. Proc. of World Academy of Science, Engineering and Technology. Vol. 21 Jan. 2007. ISSN 1307-6884.

[19] Lorenzi, F., Ricci, F. (2005). Case-based recommender systems: a unifying view. In: Intelligent Techniques in Web Personalisation. LNAI 3169. Springer-Verlag.

[20] Valtolina, S., Mussio, P., Bagnasco, G. G., Mazzoleni, P., Franzoni, S., Geroli, M., and Ridi, C. (2007). Media for knowledge creation and dissemination: semantic model and narrations for a new accessibility to cultural heritage. Proc. of the 6th ACM SIGCHI Conference on Creativity & Cognition. Washington, DC, USA, June 13 - 15, 2007. C&C '07. ACM, New York, NY, 107-116.

Video Quality Issues for Mobile Television

Carlos D. M. Regis, Daniel C. Morais
Raissa Rocha and Marcelo S. Alencar
Institute of Advanced Studies in Communications (Iecom)
Federal University of Campina Grande (UFCG)
Campina Grande, Brazil
Email: {danilo, daniel, raissa, malencar}@iecom.org.br

Mylene C. Q. Farias
Institute of Advanced Studies in Communications (Iecom)
Federal University of São Paulo (Unifesp)
São José dos Campos, Brazil
Email: mylene.farias@unifesp.br

***Abstract*—The use of mobile television requires the reduction of the image dimension, to fit on the mobile device screen. The procedure relies on space transcoding, which can be done in several ways, and this article uses down-sampling and filtering to accomplish this. Sixteen types of filter are presented to reduce the spatial video resolution from the CIF to QCIF format for use in mobile television. The objective, PSNR and SSIM, and subjective, PC, methods were used to evaluate the quality of the transcoded videos. The subjective evaluation used the H.264 encoder, with reduced bit rate and temporal resolution of the video, was implemented using a cellular device.**

***Index Terms*—Mobile television, Performance evaluation, Quality of video, Coding and processing, Transcoding.**

I. Introduction

Mobile television is a technology that allows the transmission of television programs or video to mobile devices, including cell phones and PDA's. The programs can be transmitted to a particular user in a certain area as a download process, via terrestrial broadcasting or satellite. The telecommunication operators offer video services using Digital Multimedia Broadcast (DMB), Integrated Services Digital Broadcasting Terrestrial (ISDB-T), Qualcomm MediaFLO, Digital Video Broadcasting – Handheld (DVB-H) [1], [2] and Digital Video Broadcasting – Satellite (DVB-SH) [3]. The Integrated Services Digital Broadcasting Terrestrial Built-in (ISDB-Tb) standard defines the reception of video signals in various formats for fixed or mobile receivers, with simultaneous transmission using the compression standards MPEG-2 and H.264 [4].

Table I shows a comparison of mobile television technologies based on broadcasting transmission.

TABLE I
Comparison of mobile television technologies based on broadcasting transmission.

Characteristics	DVB-H	DMB	ISDB-T
Video and audio formats	MPEG-4 or WM9 video AAC or WM audio	MPEG-4 video áudio BSAC	MPEG-4 video áudio AAC
Transport stream	IP over MPEG-2 TS	MPEG-2 TS	MPEG-2 TS
Modulation	QPSK or 16 QAM with COFDM	DQPSK with FDM	DQPSK or QPSK or 16-QAM or 64-QAM with BST-OFDM
RF bandwidth	5-8 MHz	1.54 MHz (Korea)	433 kHz (Japan)
Energy saving	Time division	Band reduction	Band reduction

In a digital television scenario the video signal may have different bit rates, encoding formats, and resolutions. Figure 1 is illustrates a block diagram of the transcoding process [5]. The video transcoder converts a video sequence to another one, including coding with different temporal and spatial resolutions and bit rates. The transcoding also saves space and production time, because only the content with maximum resolution is stored.

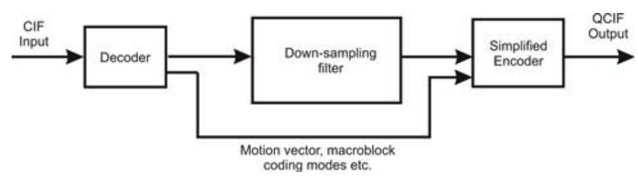


Fig. 1. The cascaded pixel domain transcoder architecture to reduce the spacial resolution.

The cell phones present several physical limitations when compared with a traditional television equipment. The main restrictions are the battery life, lower processing capacity, memory capacity and small display. Those restrictions impose limitations on the videos formats that can be played on a mobile phone or any other device for mobile reception. The length and width of the video (spatial resolution), for example, must fit the video of a small display mobile phone. If the video signal is larger than the resolution of the display, the content is not easily seen by the users.

One option is to reduce the size of the device, but this means an increase in the computational load, which is not feasible because of the limited processing ability of the mobile phones. Moreover, more processing implies an increase in energy consumption.

This paper presents a comparison among different types of spatial transcoding methods, which are intended for mobile receivers. The quality issues are discussed and a quantitative performance analysis is presented for objective and subjective video quality metrics.

II. The Transcoding Process

The transcoding process can be homogeneous, heterogeneous or use some additional functions. The homogeneous transcoding changes the bit rate and the spatial and temporal resolutions. The heterogeneous transcoding performs the conversion of standards, but also converts between the interlaced and progressive formats. The additional functions provide resistance against errors in the encoded video sequence, or add invisible or watermarks logos [6], [7]. Figure 2 represents a diagram with various transcoding function.

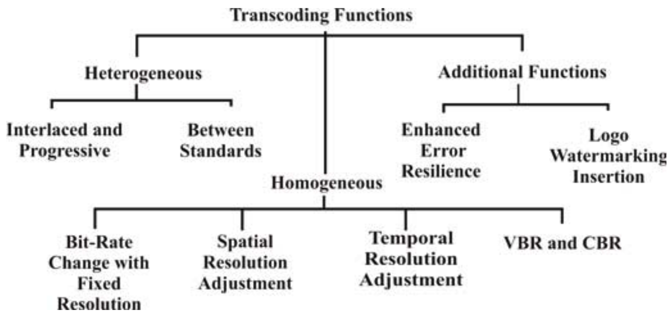


Fig. 2. Transcoding Functions.

The are two major transcoder architectures: the cascaded pixel domain transcoder (CPDT) and the DCT domain transcoder (DDT) [5]. The first one is adopted in this paper as the transcoder architecture for the CIF-to-QCIF transcoding, as shown in Figure 1. The simplified encoder is different from a stand-alone video encoder in that the motion estimation, macroblock mode decision, and some other coding processes may reuse the decoded information from the incoming video stream.

The spatial resolution reduction uses down-sampling, which changes the picture resolution from the CIF (352×288 pixels) resolution to the QCIF (176×144 pixels) format, using the down-sampling factor $352 : 176 = 2 : 1$. This factor can be achieved by up-sampling by 1 and then down-sampling by 2, as shown in Figure 3 ($S = 1$, $N = 2$), in which $h(v)$ is a low-pass filter [5].

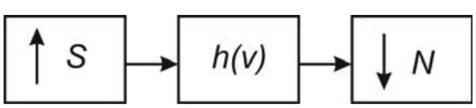


Fig. 3. The Interpolation-decimation routine for a change of M/L in terms of transmission rate.

The filters used in this article are:

- Moving Average: this technique replaces values of an $M \times M$ video block by a single pixel, which assumes the arithmetic mean of the pixels within the $M \times M$ block [8].
- Median: it provides a reorganization of the values of the pixels of an $M \times M$ block in an increasing way and chooses the central value.
- Mode: for the calculation of the mode, a comparison is made with the value that is more frequent in the $M \times M$ block [9].
- Sigma: it calculates the mean $(p(i))$ and standard deviation σ of the block $M \times M$ and verifies which pixels are within the range $(p(i)-2\sigma, p(i)+2\sigma)$. Then, the average of pixel intensities in the range is computed [10].
- Weighted Average: this technique is the average of all data entries with varying weights each, weight depends of the neighborhood pixels, as seen in Figure 4. In this case, the smoothing is less intense because there is more influence from the central pixel [11].

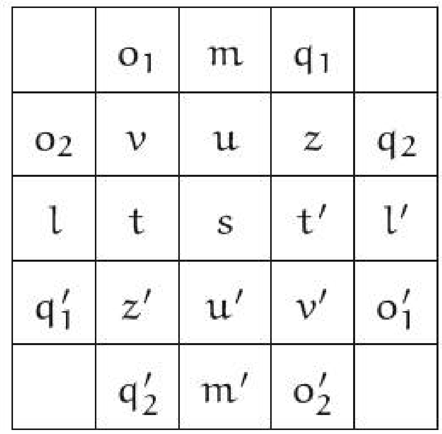


Fig. 4. Representing the neighborhood of the central pixel with value p_s.

This article presents three weighted averages, given by Equations 1, 2 and 3.

$$g(x,y) = \frac{1}{2}(x_s + \frac{1}{4}(x'_t + x_t + x'_u + x_u)) \tag{1}$$

$$g(x,y) = \frac{1}{2}(x_s + \frac{1}{5}(x'_t + x_t + x'_u + x_u + \frac{1}{4}(x'_v + x_v + x'_z + x_z))) \tag{2}$$

$$g(x,y) = \frac{1}{16}(2^3 x_s + 2^2(x'_t + x_t + x'_u + x_u) + 2^n(x'_v + x_v + x'_z + x_z)) \tag{3}$$

The transcoder used in this article includes the cited filters, with 2×2, 3×3 and 4×4 windows. For the two last ones, the videos were generated taking the pixels around the reference pixels. Those filters have been chosen for their simplicity. The moving average filter also used the 1×1 window, and was named simple elimination.

the H.264 encoder reduces the video bit rate and temporal resolution, in order to obtain the bit rates needed for the subjective tests.

III. Evaluation of the Video Transcoder

For the evaluation of a video transcoder two methods to assess the video quality are used: objective and subjective. The objective measurement is fast and simple, but there is low correlation with the human perception measurement of quality. On the other hand, the subjective measurement is expensive and time consuming.

For objective evaluation this paper uses two methods: PSNR and SSIM. The PSNR is a measure that makes the pixel to pixel comparison between the reference image and test image. The SSIM is a method that takes into account the structural information of the image, those attributes that are reflected in the structure of the objects of the scene, which depend on the average luminance and contrast of the image.

For subjective evaluation this paper is based on standard ITU-T P.910, which is the standard of subjective evaluation for multimedia [12]. The standard is mentioned three forms of assessment: Absolute Category Rating (ACR), Degradation Category Rating (DCR) and Pair Comparison (PC). The method used in this paper uses the PC method.

A. SSIM

The structural similarity metric (SSIM) is attracting the attention of the research community because of the good results obtained in the perceived quality of representation [13]. The SSIM measures how the video structure differs from the structure of the reference video, involving the evaluation of the structural similarity of the video.

The SSIM indexing algorithm is used for quality assessment of still images, with a sliding window approach. The window size 8×8 is used in this paper. The SSIM metrics define the luminance, contrast and structure comparison measures, as defined in Equation 4 [14], [15].

$$l(x,y) = \frac{2\mu_x\mu_y}{\mu_x^2+\mu_y^2}, \quad c(x,y) = \frac{2\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2}, \quad s(x,y) = \frac{\sigma_{xy}}{\sigma_x\sigma_y}, \tag{4}$$

and SSIM metrics are given in Equation 5

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)}. \tag{5}$$

The constants, C_1 and C_2, are defined in Equation 6

$$C_1 = (K_1L)^2 \quad \text{and} \quad C_2 = (K_2L)^2, \tag{6}$$

in which L is the dynamic range of the pixel values, and K_1 and K_2 are two constants whose values must be small, such that C_1 or C_2 will cause effect only when $(\mu_x^2+\mu_y^2)$ or $(\sigma_x^2+\sigma_y^2)$ is small. For all experiments in this paper, one sets $K_1 = 0.01$ and $K_2 = 0.03$, respectively, and $L = 255$, for 8 bits/pixel gray scale images. The quality measure of a video is between 0 and 1, with 1 as the best value.

B. PC

This method was chosen because the test sequences are presented in pairs, making a better comparison between the methods of transcoding.

The PC method consists of test systems (A, B, C, etc.) that are arranged in all possible $n(n-1)$ combinations of type AB, BA, CA, etc.. Thus, all pairs are displayed in both possible orders (eg AB, BA). After each pair presentation, the subject decides which video has the best quality.

The method specify that, after each presentation, the participants are invited to assess the quality of the indicated sequence. The average time for the presentation and vote should be equal or less than 10 s, depending on engine voting process used. The presentation time may be reduced or increased, according to content.

Tests were carried out with 20 people. Each participant watched six video four times generating 120 samples per video. The participants marked the quality score of a video clip on an answer sheet using a discrete scale from 0 up to 10.

A cell phone (NOKIA N95) was used for the field tests. The distance between the participants and the device is 18 cm. This distance is calculated by multiplying the smaller device screen dimension by six (3×6 cm). The tests lasted an average of 30 minutes.

IV. RESULTS

This section presents the results of the transcoded videos and the same video transcoded after coding, then the comparison is made.

For analysis of the videos was used the Mobile, News and Foreman videos [16], with 10 s for each one. These videos were chosen for displaying the following characteristics:

- Mobile: high texture and slow movement, Figure 5;
- News: little texture and slow movement, Figure 6;
- Foreman: reasonable texture and rapid movement, Figure 7.

Fig. 5. Mobile Video.

Fig. 6. News Video.

Fig. 7. Foreman Video.

A. Objective Evaluation

The efficiency of a transcoder is evaluated by the PSNR and SSIM for the processed videos. Table II and Figure 8 show the result and the PSNR curves, respectively, for transcoded videos.

TABLE II
PSNR RESULTS.

Numeber	Filter	Mobile	News	Foreman
1	Simple Elimination	37,93	45,87	48,69
2	2×2 Moving Average	39,19	42,46	44,13
3	3×3 Moving Average	41,35	48,76	52,79
4	4×4 Moving Average	40,71	45,20	48,51
5	2×2 Median	42,55	52,39	51,94
6	3×3 Median	40,22	47,44	51,51
7	4×4 Median	42,67	50,66	51,93
8	2×2 Mode	39,62	44,82	45,38
9	3×3 Mode	36,70	42,61	45,34
10	4×4 Mode	35,06	42,06	44,97
11	Weighted Average 1	39,76	47,71	51,80
12	Weighted Average 2	39,96	47,93	52,07
13	Weighted Average 3	40,91	48,69	53,02
14	2×2 Sigma	43,10	53,33	52,62
15	3×3 Sigma	41,12	48,65	52,62
16	4×4 Sigma	43,12	50,03	51,59

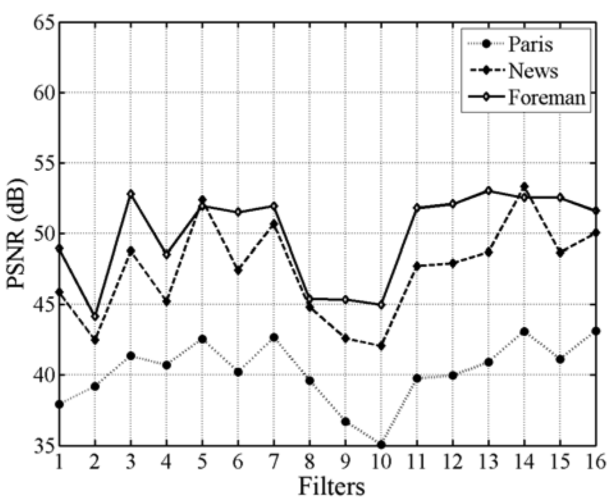


Fig. 8. PSNR curves for the transcoded videos.

For the Mobile video the test showed that the best results were processed with the 4×4 Sigma, 2×2 Sigma and 4×4 Median filter. News video the best results correspond the videos processed with the 2×2 Sigma, 2×2 Median and 4×4 Median filter. For the Foreman video the best results correspond to the processed video with Weighted Average 3, 3×3 Moving Average and 2×2 sigma filter. These results are shown in Figure 8.

The Table III and Figure 9 show the result and the SSIM curves, respectively, for transcoded videos.

TABLE III
SSIM RESULTS.

Numeber	Filter	Mobile	News	Foreman
1	Simple Elimination	0,9713	0,9793	0,97835
2	2×2 Moving Average	0,9842	0,9618	0,9739
3	3×3 Moving Average	0,9785	0,9744	0,9841
4	4×4 Moving Average	0,9511	0,9560	0,9701
5	2×2 Median	0,9910	0,9901	0,9942
6	3×3 Median	0,9828	0,9805	0,9882
7	4×4 Median	0,9578	0,9704	0,9799
8	2×2 Mode	0,9765	0,9752	0,9764
9	3×3 Mode	0,9707	0,9705	0,9726
10	4×4 Mode	0,9647	0,9732	0,9725
11	Weighted Average 1	0,9871	0,9819	0,9887
12	Weighted Average 2	0,9877	0,9814	0,9882
13	Weighted Average 3	0,9859	0,9786	0,9867
14	2×2 Sigma	0,9918	0,9885	0,9939
15	3×3 Sigma	0,9813	0,9783	0,9860
16	4×4 Sigma	0,9593	0,9636	0,9751

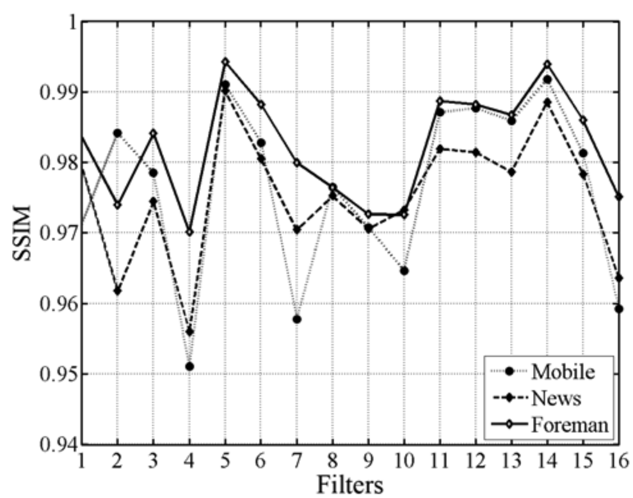


Fig. 9. SSIM curves for the transcoded videos.

It can be observed from Table III that the best results for the Mobile video were obtained using the 2×2 Sigma, 2×2 Median and Weighted Average 2 filters. In the videos News and Foreman the best results using the 2×2 Median, 2×2 Sigma and Weighted Average 1 filter.

With the results of PSNR and SSIM methods could find the correlation between the measures. The correlation in the Mobile video was obtained 0.1408, which is a weak correlation well. For the News video to get the measures correlation found was 0.5424, which is a correlation average. For the Foreman video the correlation between measures obtained was 0.7492, which is a strong correlation.

B. Processing Time

Regarding the processing time, it is possible to analyze the increase in time as the filter window increases, as shown in Table IV. This table shows that the sigma and mode filters demand longer processing periods as compared with the moving average and the weighted average filters, and the median processing time is slightly higher than the average.

The best results considering the processing time was the simple elimination, weighted average, 2×2 and 3×3 moving average and the 2×2 median.

The results for the sigma filter shown in the Table IV are given as the average of the obtained values, because each window is related to the pixel number.

TABLE IV
PROCESSING TIME FOR A VIDEO

Transcoding Method	Time(seconds)
Simple Elimination	0.47
2×2 Moving Average	1.30
3×3 Moving Average	1.13
4×4 Moving Average	3.89
2×2 Median	1.59
3×3 Median	5.00
4×4 Median	13.69
2×2 Mode	7.78
3×3 Mode	8.22
4×4 Mode	56.47
Weighted Average 1	0.75
Weighted Average 2	1.19
Weighted Average 3	3.42
2×2 Sigma	5.76
3×3 Sigma	12.06
4×4 Sigma	20.50

C. Subjective Evaluation

The evaluation of the transcoder with the subjective method used eight videos, that were transcoded using the Weighted Average 1, 2×2 Sigma, 2×2 Median, Weighted Average 2, 3×3 Sigma, 3×3 Median, 3×3 Moving Average and Weighted Average 3 filter.

The subjective tests were performed using the PC, the device N95 and all the videos encoded using H.264 encoder with bit rate of 243 kbit/s and 15 frames/s.

For the Foreman video the values of MOS are shown in Table V and Figure 10. The best result for this video the was transcoded video using the 2×2 Sigma, 2×2 Median, Weighted Average 3 and 3×3 Median filter.

Mobile video for the values of MOS are shown in Table VI and Figure 11. The best results for the videos that video was transcoded using the Weighted Average 3 and 3×3 Median filter .

Video News for the values of MOS are shown in Table VII and Figure 12. The best results for the videos that video was transcoded using the 2×2 Sigma and 2×2 Median filter.

TABLE V
MOS FOREMAN VIDEO.

Videos	Nome	MOS
1	2×2 Sigma	7.7899
2	2×2 Median	7.5630
3	3×3 Moving Average	7.3644
4	Weighted Average 3	7.5000
5	3×3 Sigma	7.3083
6	Weighted Average 2	7.2857
7	Weighted Average 1	7.2250
8	3×3 Median	7.5000

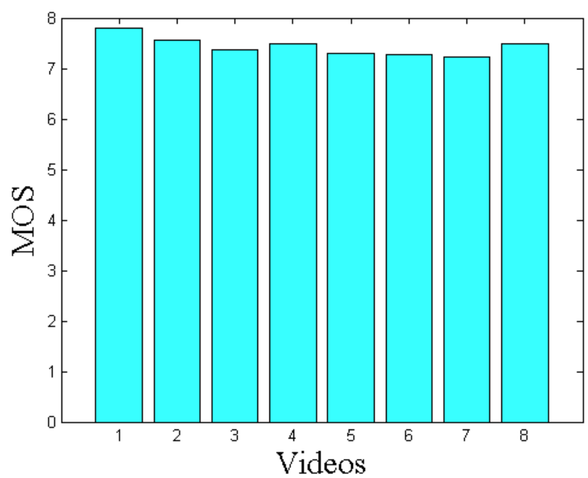


Fig. 10. MOS Foreman video.

TABLE VI
MOS MOBILE VIDEO.

Videos	Nome	MOS
1	2×2 Sigma	7.3167
2	2×2 Median	7.3025
3	3×3 Moving Average	7.4417
4	Weighted Average 3	7.5630
5	3×3 Sigma	7.1849
6	Weighted Average 2	7.1513
7	Weighted Average 1	7.2083
8	3×3 Median	7.5042

TABLE VII
MOS MOBILE VIDEO.

Videos	Nome	MOS
1	2×2 Sigma	7.7500
2	2×2 Median	7.4538
3	3×3 Moving Average	7.0167
4	Weighted Average 3	7.3140
5	3×3 Sigma	7.2917
6	Weighted Average 2	7.3250
7	Weighted Average 1	7.3109
8	3×3 Median	7.2231

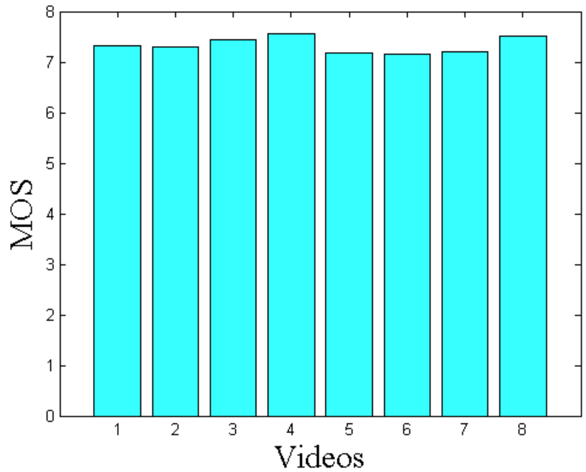


Fig. 11. MOS Mobile video.

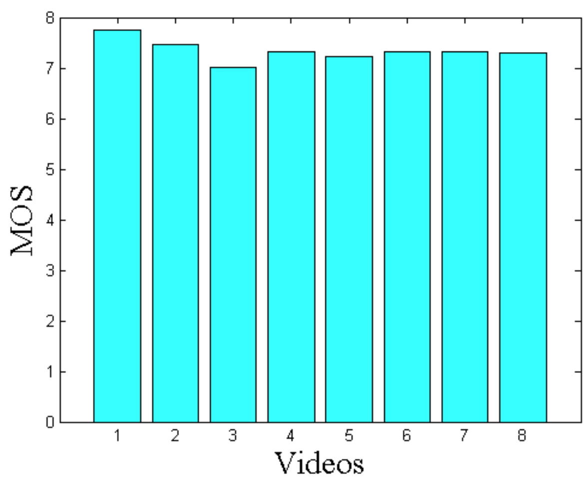


Fig. 12. MOS News video.

The correlation between the MOS and PSNR results for each video was calculated, resulting in a low correlation to the videos Foreman and Mobile, 0.3721 and 0.3209, respectively, and a strong correlation to the video News, 0.7745.

Already a correlation between the MOS and SSIM values obtained better, as expected. For the Foreman video the correlation between the SSIM and the MOS is average, 0.5837, for the Mobile video the correlation is weak, - 0.372 and Video News the correlation is strong, 0.8486.

The filters that have the best results were the 2×2 Sigma, 3×3 Median, Weighted Average 3 and 2×2 Median.

V. Conclusion

The article discussed the characteristics of mobile television, mainly related to quality issues. It has been shown that the spatially transcoded videos for this service presented satisfactory results, since all results provided acceptable PSNRs. For the evaluation using the PSNR method the 4×4 median, 2×2 Sigma and 2×2 median filter produced the best result. For the SSIM method the 2×2 Sigma and 2×2 Median showed the best results.For the subjective tests, the spatial transcoded videos using 2×2 Median and 2×2 Sigma filters obtained better results.

As the spatially transcoded videos using 2×2 Median and 2×2 Sigma filters give the best results for both objective and subjective measures, one concludes that the techniques are appropriate to space transcoding. The 2×2 Median has a small advantage over the 2×2 Sigma regarding the required time for processing.

The correlation results show that the SSIM method presents a better correlation with the subjective tests, when compared with the PSNR method. Depending on the video technique the SSIM presents a low correlation with the subjective tests.

VI. Acknowledgement

The authors acknowledge the financial support from Capes and CNPq, and thank the Iecom for using its structure and equipment.

References

[1] D. T. T. A. Group, "Television on a handheld receiver, broadcasting with DVB-H," Geneva, Switzerland, 2005.

[2] A. Kumar, *Mobile TV: DVB-H, DMB, 3G Systems and Rich Media Applications*. Focal Press Media tecnology Professional, 2007.

[3] D. V. Broadcasting, "DVB approves DVB-SH specification - new specification addresses delivery of multimedia services to hybrid satellite/terrestrial mobile devices," 2007.

[4] M. S. Alencar, *Digital Television Systems*. New York: Cambridge University Press, 2009.

[5] J. Xin, M.-T. Sun, B.-S. Choi, and K.-W. Chun, "An HDTV-to-SDTV spatial transcoder," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 12, no. 11, pp. 998–1008, Nov 2002.

[6] I. Ahmad, X. Wei, Y. Sun, and Y.-Q. Zhang, "Video transcoding: an overview of various techniques and research issues," *Multimedia, IEEE Transactions on*, vol. 7, no. 5, pp. 793–804, Oct. 2005.

[7] J. Xin, C.-W. Lin, and M.-T. Sun, "Digital video transcoding," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 84–97, Jan. 2005.

[8] T. Acharya and A. K. Ray, *Image Processing - Principles and Applications*. Hoboken, New Jersey, USA: John Wiley & Sons, Inc., 2005.

[9] H. Wu and K. Rao, *Digital Video Image Quality and Perceptual Coding*. Boca Raton, FL, USA: CRC Press Taylor & Francis Group, 2006.

[10] R. Lukac, B. Smolka, K. Plataniotis, and A. Venetsanopoulos, "Generalized adaptive vector sigma filters," *International Conference on Multimedia and Expo. ICME '03.*, vol. 1, pp. I–537–40 vol.1, July 2003.

[11] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 2001.

[12] ITU-T, "ITU-T recommendation P.910, subjective video quality assessment methods for multimedia applications," September 1999.

[13] R. de Freitas Zampolo, D. de Azevedo Gomes, and R. Seara, "Avaliação e comparação de métricas de referência completa na caracterização de limiares de detecção em imagens," *XXVI Simpósio Brasileiro de Telecomunicações - SBrT 2008*, Sept. 2008.

[14] L. L. Zhou Wang and A. C. Bovik, "Video quality assessment based on structural distortion measurement," *Signal Processing: Image Communication*, vol. 19, no. 2, pp. 121–132, february 2004.

[15] M. Vranjes, S. Rimac-Drlje, and D. Zagar, "Objective video quality metrics," *49th International Symposium ELMAR, 2007*, pp. 45–49, Sept. 2007.

[16] "Yuv video sequences," http://trace.eas.asu.edu/yuv/index.html, November 2008.

Comparing the "Eco Controllo" 's video *codec* with respect to MPEG4 and H264

Claudio Cappelli[1]

[1] Eco Controllo SpA
Via Camillo De Nardis 10, 80127 Napoli (NA), Italy
claudio.cappelli@ecocontrollo.com

Abstract

This paper reports results of an experimental comparison between the video codec produced by the company Eco Controllo SpA and those of main commercial standards, such as MPEG-4 and H.264. In particular, the experiments aimed to test the ratio between the quality of the compressed image and the achieved bit rate, where the quality of the compressed image is meant as high or low fidelity with respect to the original image. Such fidelity has been measured by means of both objective and subjective tests. In particular, as for the formers, the Peak Signal to Noise Ratio (PSNR) and the Structural Similarity Measure (SSIM) have been used. As for subjective tests, the Double Stimulus Impairment Scale (DSIS) methodology standardized by International Telecommunication Union (ITU) has been employed [1]. The tests have been repeated for different video resolutions (corresponding to the different video formats PAL and HD), different frame rates (25, 30, etc.), and different values of bit rate. Finally, it has been beyond these experiments the evaluation of critical aspects concerning live video transmission.

1 Introduction

Digital images are subject to several distortions introduced during acquisition, processing, compression, storage, transmission, and reproduction phases, each of which can decrease the vision quality. Since images are intended for human beings, the natural way to quantify their quality is to use subjective evaluation.

The methodologies for subjective analysis have been standardized by *International Telecommunication Union* (*ITU*) [1], aiming to make such tests reproducible and verifiable. In practice, subjective tests consist in presenting a selection of images and videos to a sample of the population. Users watch video contents and express a vote based on the perceived quality, highlighting the presence of aberrations, or distortions, with respect to a given content of reference. The results are opportunely elaborated, and enable the evaluation of the average quality of the system under examination.

Objective Quality Metrics represent an alternative to subjective metrics. They allow us to considerably reduce costs, since the test they prescribe can be accomplished much more rapidly. *Objective Quality Metrics* derive from *subjective analysis*, representing a kind of abstraction or theoretical model of them. They can be classified based on the presence or absence of a reference system (an original video or image without distortions), which the system under examination can be compared with. Many existing comparison systems are considered "full-reference", meaning that every system under evaluation can be compared with a reference system without distortions. Nevertheless, in many practical situations it is not possible to use a reference system, and in such cases it is necessary to adopt a so called "no-reference" or "blind" approach. A third situation is that in which there is a partial availability of a reference system, that is, only some basic characteristics of the reference system are known. In such a case, the available information can be considered as a valid support for evaluating the quality of the system under examination. This approach is referred to as "reduced-reference".

The simplest and most widely adopted "full-reference" metric is the so called "peak signal-to-noise ratio" (*PSNR*), based on the *mean square error* (*MSE*), which is in turn computed by averaging squares of differences in intensity between homologous pixels of the compressed and the reference images. *PSNR* is simple to compute and it has a clear meaning. Nevertheless, it does not always reflect the visual quality as it is perceived by humans [3, 4, 5, 6, 7, 8, 9, 10, 11].

In the last three decades, a considerable effort has been made to develop objective quality metrics exploiting the known characteristics of the Human Vision System (HVS). An example of such metrics is the *SSIM* index: Structural Similarity Measure. *SSIM* index compares patterns of pixels based on the intensity normalized with respect with luminosity and contrast.

This paper describes the results of an experimental comparison between the video codec produced by the company *Eco Controllo SpA* and two main standards, such as *MPEG-4* and *H.264*. *Eco Controllo* has commissioned such comparison to the italian research center on ICT *Cerict*, which has accomplished them by means of both objective and subjective metrics. For the objective analysis they have used both *PSNR* and *SSIM* index. Regarding the subjective analysis they have used the *DSIS* technique [1].

The paper is organized as follows, Section 2 describes the type of tests that have been performed, including test parameters and characteristics of hardware used, Section 3 describes the test cases used, Section 4 describes results of objective tests, and Section 4 those of subjective tests. Finally, conclusions are given in Section 6.

2 Comparative Tests

The tests have been of type *Full-Reference*, and have produced both objective and subjective analysis, aiming to evaluate compression quality. The comparative study has been executed on a sample of files compressed in *batch* modality, that is, first all the original files have been compressed, and then they have been analyzed. The only constraints codecs had to abide by were the compliance with required *bit rate*, and the size of the compressed file. Although *Eco Controllo SpA* aims to use its codec for live broadcasting, it has been beyond the scope of this test the evaluation of possible critical issues arising during transmission and reception of video signals, and issues related to the hardware and software resources needed to execute the selected codecs. Furthermore, no constraints have been imposed on time needed to compress videos.

The Codec produced by *Eco Controllo SpA* has been compared with main known standards, such as *MPEG-4* e *H.264*. To this end, it has been chosen a unique commercial software embedding both these codecs. All the compared codecs have been tested by using their respective default parameters, and without human intervention. In particular, the *Simple Profile* has been used for *MPEG-4* compression, and the *Main Profile* for *H.264* compression. Moreover, beyond the specification of the *bit rate* and the *frame rate*, no other parameters have been specified, and no pre/post-production work has been performed. Finally, compressions and tests have been executed on a *Siemens Celsius V830 Workstation*, whose characteristics are described in table 1.

Table 1: Workstation Siemens Celsius V830

RAM	8 GB
CPU	2 AMD Opteron 240
HD	2 HD SataII di 400GB
VID	NVIDIA Quadro FX 3400 - 256 MB
OS	Windows XP64

3 Test Cases

When executing comparative tests it is particularly important to choose a significative test set. Using a standard test set has the advantage of providing comparable test results, often reducing the cost of tests. On the other hand, exclusively using well known video sequences potentially reduces test integrity, since it cannot be prevented the use of ad-hoc compression techniques, optimized for public available test sets. In this experimental comparison several types of test sets have been used, including test sets commonly used in scientific studies of this area, and heterogeneous video sequences commonly used in television programs, hence realized with professional quality. In particular, test cases have been selected among the following video test sets:

- HDTV (720p - 50Hz e 25 Hz) "SVT High Definition Multi Format Test Set" [12] - Video sequences produced by the swedish television channel *SVT*, also available on the 'Video Quality Experts Group (VQEG)' web site `ftp://vqeg.its.bldrdoc.gov/HDTV/SVT_MultiFormat/`. Moreover, the same video sequences have been reduced to derive the *PAL* video sequences used in the tests.
- CIF - Test Set: Xiph.org. In particular the "Derf" collection, available at `http://media.xiph.org`.

In order to execute tests, a sample of 17 videos in the three different formats (*720P@25Hz*, *720P@50Hz*, *PAL@25Hz*) has been selected. Such formats have been chosen to test compression algorithms with respect to the standards that are currently, and in the near future, used in the television field. In particular, the choice of the *PAL* format has been done to test our system with respect to the technology currently used in television transmission systems, whereas the *720P* resolution will be the one used in the near future with the introduction of the so called "high resolution".

4 Objective Tests on Video 720P and *PAL*

Objective tests have been executed by using *PSNR* and *SSIM* metrics on a database of *PAL* and *FullHD* videos.

Three series of 17 video sequences, in *PAL@25Hz*, *720P@25Hz*, and *720P@50Hz* formats, respectively, have been compressed by using the video codecs *H.264*, *MPEG-4*, and the video codec by *Eco Controllo*. Each video sequence has been compressed at 500, 1000, 2000, 3000 e

4000Kbps, yielding 765 different compressed files. Among these, only those having size $\pm 5\%$ di F have been considered, where,

$$F = \frac{br \cdot s}{8} \tag{1}$$

where:

br :	*bit rate* per second in $Kbps$ (1000bit/sec)
s :	video duration in seconds
F :	file size in $KBytes$

Among the 765 produced file, only 507 resulted valid after compression with the requested *bit rate*, and have successively been evaluated through the *PSNR* and *SSIM* metrics, by using the *MSU Video Quality Measurement Tool* rel. 1.4, produced by the *Graphics & Media Lab Video Group* of *Moscow State University*.

Table 2: Test results through the *SSIM* metrics

	EcoControllo	H264	MPG4
$\overline{u}$	0.90	0.85	0.78
S	0.06	0.09	0.17
δ	0.01	0.02	0.03
ci	[0.89,0.91]	[0.83,0.87]	[0.75,0.81]
min	0.71	0.58	0.00
max	0.99	0.99	0.97

The results, synthesized in tables 2 and 3, reveal that the codec *Eco Controllo* has preserved the best quality with respect to the two selected metrics, both on the average and on each analyzed video sequence. Moreover, the *Eco Controllo* codec has turned out to be more stable with respect to the tested video sequences, that is, the gap among single test sessions is lower than the one observed with the codecs *H.264*, and *MPEG-4*, respectively. This is confirmed by the confidence interval and by figures 5, 1, 6, 2, 4, and 3.

Another interesting characteristics to notice is that the *Eco Controllo* and *H.264* codecs reach the same average maximum vote, and the same can be said for the *MPEG-4* codec. Nevertheless, by observing worst cases, it can be noticed that sometimes the *MPEG-4* codec fails with the requested *bit rate*, which in the figure is indicated with a 0 value, whereas the *H.264* codec shows *SSIM* a poor value, and the *Eco Controllo* codec keeps a relatively high score, never going below the average $SSIM$ = *0.71*.

Table 3: Test results through the *PSNR* metrics

	EcoControllo	H264	MPG4
$\overline{u}$	35.08	31.79	29.48
S	4.04	4.44	3.07
δ	0.68	0.75	0.57
ci	[34.4,35.76]	[31.04,32.54]	[28.91,30.05]
min	26.72	21.83	22.91
max	46.00	43.58	37.73

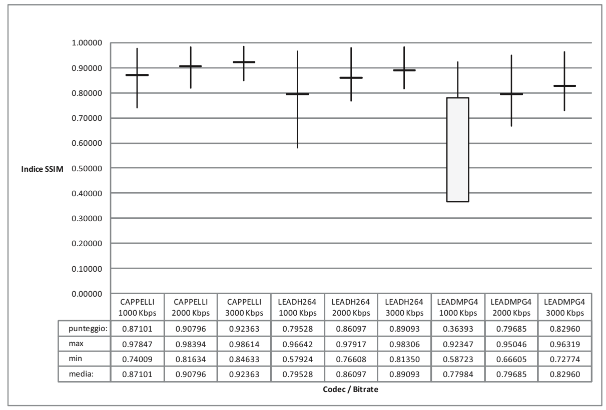


Figure 1: Comparison with the *SSIM* metric - 720P a 25Hz

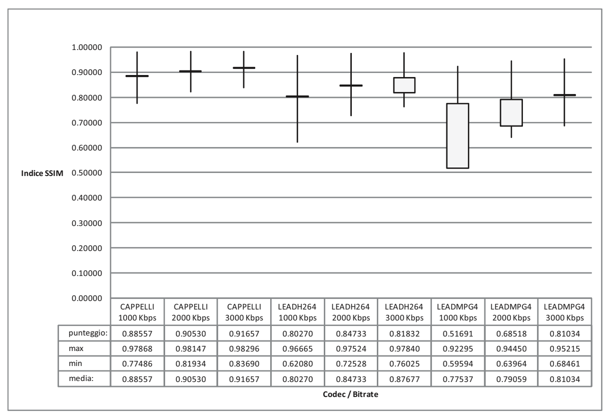
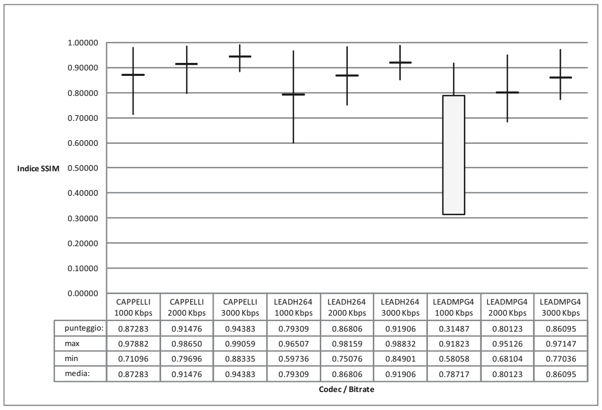


Figure 2: Comparison with the *SSIM* metric - 720P a 50Hz

5 Subjective Tests on Video 720P and *PAL*

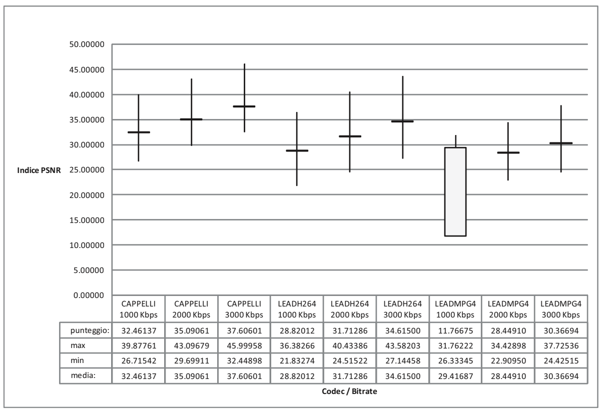
In order to validate results of objective tests, the selected codecs have been further compared through subjective tests accomplished by means of the *DSIS* method [1]. In particular, 8 video sequences have been randomly selected, and successively shown at three different *bit rates* (1000,2000 e 3000Kbps) to 16 human evaluators. These have been subdivided in two different groups, each participating to a different evaluation session of 30 minutes.

User evaluation data available on paper support have been digitized and successively processed according to the *DSIS* methodology, yielding the results reported in tables 4 e 5.



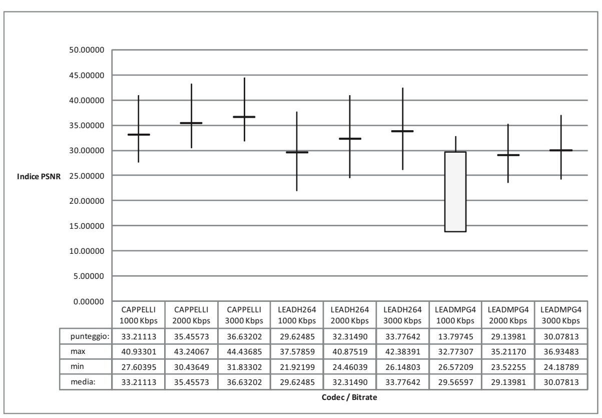
	CAPPELLI 1000 Kbps	CAPPELLI 2000 Kbps	CAPPELLI 3000 Kbps	LEADH264 1000 Kbps	LEADH264 2000 Kbps	LEADH264 3000 Kbps	LEADMPG4 1000 Kbps	LEADMPG4 2000 Kbps	LEADMPG4 3000 Kbps
punteggio:	0.87283	0.91476	0.94383	0.79309	0.86806	0.91906	0.31487	0.80123	0.86095
max	0.97882	0.98650	0.99059	0.96507	0.98159	0.98832	0.91823	0.95126	0.97147
min	0.71096	0.79696	0.88335	0.59736	0.75076	0.84901	0.58058	0.68104	0.77036
media:	0.87283	0.91476	0.94383	0.79309	0.86806	0.91906	0.78717	0.80123	0.86095

Figure 3: Comparison with the *SSIM* metric - PAL a 25Hz



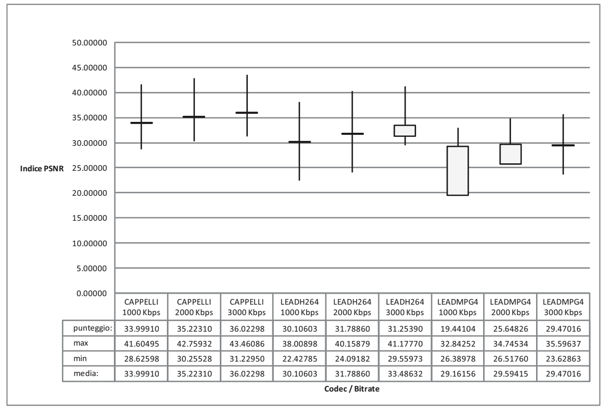
	CAPPELLI 1000 Kbps	CAPPELLI 2000 Kbps	CAPPELLI 3000 Kbps	LEADH264 1000 Kbps	LEADH264 2000 Kbps	LEADH264 3000 Kbps	LEADMPG4 1000 Kbps	LEADMPG4 2000 Kbps	LEADMPG4 3000 Kbps
punteggio:	32.46137	35.09061	37.60601	28.82012	31.71286	34.61500	11.76675	28.44910	30.36694
max	39.87761	43.09679	45.99958	36.38266	40.43386	43.58203	31.76222	34.42898	37.72536
min	26.71542	29.69911	32.44898	21.83274	24.51522	27.14458	26.33345	22.90950	24.42515
media:	32.46137	35.09061	37.60601	28.82012	31.71286	34.61500	29.41687	28.44910	30.36694

Figure 4: Comparison with the *PSNR* metric - PAL a 25Hz



	CAPPELLI 1000 Kbps	CAPPELLI 2000 Kbps	CAPPELLI 3000 Kbps	LEADH264 1000 Kbps	LEADH264 2000 Kbps	LEADH264 3000 Kbps	LEADMPG4 1000 Kbps	LEADMPG4 2000 Kbps	LEADMPG4 3000 Kbps
punteggio:	33.21113	35.45573	36.63202	29.62485	32.31490	33.77642	13.79745	29.13981	30.07813
max	40.93301	43.24067	44.43685	37.57859	40.87519	42.38391	32.77307	35.21170	36.93483
min	27.60395	30.43649	31.83302	21.92199	24.46039	26.14803	26.57209	23.52255	24.18789
media:	33.21113	35.45573	36.63202	29.62485	32.31490	33.77642	29.56597	29.13981	30.07813

Figure 5: Comparison with the *PSNR* metric - 720P a 25Hz



	CAPPELLI 1000 Kbps	CAPPELLI 2000 Kbps	CAPPELLI 3000 Kbps	LEADH264 1000 Kbps	LEADH264 2000 Kbps	LEADH264 3000 Kbps	LEADMPG4 1000 Kbps	LEADMPG4 2000 Kbps	LEADMPG4 3000 Kbps
punteggio:	33.99910	35.22310	36.02298	30.10603	31.78860	31.25390	19.44104	25.64826	29.47016
max	41.60495	42.75932	43.46086	38.00898	40.15879	41.17770	32.84252	34.74534	35.59637
min	28.62598	30.25528	31.22950	22.42785	24.09182	29.55973	26.38978	26.51760	23.62863
media:	33.99910	35.22310	36.02298	30.10603	31.78860	33.48632	29.16156	29.59415	29.47016

Figure 6: Comparison with the *PSNR* metric - 720P a 50Hz

Only video in 720P (e.g. 1280x720 pixels) format have been selected, with both 25 and 50 Hz video frame frequencies. The choice of such parameters is motivated by the fact that they are used within the Digital Television and High Definition Digital TV (HDTV) in all those countries (Italy included) traditionally using the *PAL* and *SECAM* video transmission systems.

In order to derive more meaningful results, low *bit rates* have been used to stress the selected codecs and test their behavior under critical conditions.

As prescribed by the *DSIS* methodology, evaluators have been placed in a comfortable room, and seated in positions guaranteeing an appropriate visualization angle with respect to a *FullHD* plasma monitor used to show video sequences. Evaluators have been requested to express the quality of each shown video sequence by choosing one of the following options:

- no defect
- visible but not noisy defects
- slightly noisy defects
- noisy defects
- highly noisy defects

Evaluators have been selected among students and workers. Each of them has preventively undergone the Ishihara test for color blindness. The latter is a test published by Prof. *Shinobu Ishihara* in 1917, and it consists in submitting to the user several colored disks, named Ishihara disks, each containing a circle of colored points arranged to form a number visible to people without color blindness problems, and invisible for people having some problems to this regard, especially in the perception of red and green colors [2].

Table 4: Subjective Analysis results

	Eco Controllo	H264	MPG4
$\overline{u}$	4.76	3.23	2.09
S	0.47	1.28	0.97
δ	0.08	0.22	0.19
ci	[4.68,4.84]	[3,3.45]	[1.9,2.29]
$\overline{ci}$	[0.94,0.97]	[0.60,0.69]	[0.38,0.46]
$Score$	0.95	0.65	0.42

Table 5: Subjective Analysis results, grouped by Bit rate

	Eco Controllo		
Bit Rate	1000KBps	2000KBps	3000KBps
$\overline{u}$	4.78	4.72	4.81
S	0.42	0.52	0.40
δ	0.15	0.13	0.14
ci	**[4.64,4.93]**	**[4.59,4.85]**	**[4.68,4.95]**
$score$	**0.96**	**0.94**	**0.96**
	H264		
Bit Rate	1000KBps	2000KBps	3000KBps
$\overline{u}$	2.97	3.34	3.25
S	1.06	1.29	1.46
δ	0.37	0.32	0.51
ci	**[2.6,3.34]**	**[3.03,3.66]**	**[2.74,3.76]**
$score$	**0.59**	**0.67**	**0.65**
	MPG4		
Bit Rate	1000KBps	2000KBps	3000KBps
$\overline{u}$	1.50	2.33	2.21
S	0.78	1.10	0.59
δ	0.31	0.31	0.24
ci	**[1.19,1.81]**	**[2.02,2.64]**	**[1.97,2.44]**
$score$	**0.30**	**0.47**	**0.44**

Such results essentially confirm those derived with objective metrics, even though the gap among different codecs here is amplified. In particular, the support for the confidence index shown in tables 4 and 5, seems to highlight a greater stability of the *Eco Controllo's* algorithm. Even in this case, by considering the maximum average score, the algorithms *H.264* and *Eco Controllo* achieve similar results, which probably means that users do not perceive meaningful defects when codecs are used with less demanding bit rates. Nevertheless, in the worst and average case the *Eco Controllo's* codec achieves more precise scores, that is, with less variations with respect to other codecs. Thus, under the test conditions described here, the *Eco Controllo's* codecs showed better performances with respect to the other selected codecs.

6 Conclusions

The successful diffusion of digital video applications depends on the capability to have low cost transmission systems for high quality video sequences. This means to be able to achieve high compression ratios of images in order to transmit them on low bandwidth networks, yielding considerable cost reductions. However, in doing this it is necessary to preserve adequate quality of compressed images with respect to original images. This work described the results of an experimental comparison of the video codecs produces by Eco controllo with respect to main commercial standards, by using several test methodologies described in the literature, and a conspicuous number of heterogeneous video sequences. According to results of such tests, both objective and subjective test methodologies described in this paper have revealed a better quality of video sequences compressed through Eco controllo's codecs, for each chosen bit rate.

References

[1] ITU-R. 'Methodology for the subjective assessment of the quality of television pictures', RECOMMENDATION ITU-R BT.500-11, 1–48, 2002. 1, 2, 3

[2] S. Ishihara. 'Tests for colour-blindness', Handaya, Tokyo, Hongo Harukicho, 1917. 4

[3] B. Girod, 'What's wrong with mean-squared error', in Digital Images and Human Vision, A. B. Watson, Ed. Cambridge, MA: MIT Press, pp.207-220, 1993. 1

[4] P. C. Teo and D. J. Heeger, 'Perceptual image distortion', in Proc. SPIE, vol. 2179, pp. 127-141, 1994. 1

[5] A. M. Eskicioglu and P. S. Fisher, 'Image quality measures and their performance', IEEE Trans. Commun., vol. 43, pp. 2959-2965, Dec. 1995. 1

[6] M. P. Eckert and A. P. Bradley, ‘Perceptual quality metrics applied to still image compression’, Signal Processing, vol. 70, pp. 177-200, Nov. 1998. 1

[7] S. Winkler, ‘A perceptual distortion metric for digital color video’, in Proc. SPIE, vol. 3644, pp. 175-184, 1999. 1

[8] Z. Wang, ‘Rate scalable Foveated image and video communications’, Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. Texas at Austin, Austin, TX, Dec. 2001. 1

[9] Z. Wang and A. C. Bovik, ‘A universal image quality index’, IEEE Signal Processing Letters, vol. 9, pp. 81-84, Mar. 2002. 1

[10] Z. Wang. ‘Demo Images and Free Software for “a Universal Image Quality Index”’. Available: http://anchovy.ece.utexas.edu/~zwang/research/quality_index/demo.html 1

[11] Z. Wang, A. C. Bovik, and L. Lu, ‘Why is image quality assessment so difficult’, in Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, vol. 4, Orlando, FL, pp. 3313-3316, May 2002. 1

[12] L. Haglund, ‘The SVT High Definition Multi Format Test Set’, Sveriges Television, 2006. Available: ftp://vqeg.its.bldrdoc.gov/HDTV/SVT_MultiFormat/SVT_MultiFormat_v10.pdf

2

An Experimental Evaluation of the Mobile Channel Performance of the Brazilian Digital Television System

Carlos D. M. Regis and Marcelo S. Alencar
Institute of Advanced Studies in Communications (Iecom)
Federal University of Campina Grande (UFCG)
Campina Grande, Brazil
Email: {danilo, malencar}@iecom.org.br

Jean Felipe F. de Oliveira
Positivo Informática S/A
Curitiba, Brazil
Email: jeanfelipe@positivo.com.br

***Abstract*—This work presents an analysis of the mobile channel of the Brazilian Digital Television System. With the advent of this system, diverse conditions must be emphasized, which pose an impact on the development of the transmission equipment. The key variables that influence the degradation of the quality of the digital signal are the mobile television velocity, the number of fading components, the random phase shift, the propagation delay and the Doppler Effect. A robust knowledge about the behavior of those variables is important to evaluate the channel transmission, and to design the equipment in accordance with the available standards. Based on the study of the impact of those factors a separate assessment of the influence of each variable in the quality of the demodulated constellations is proposed, and its relevance on the transmission process. This research was conducted at Positivo Informática S/A digital TV laboratory.**

***Index Terms*—Mobile Television, ISDB-Tb, ISDTV, Digital TV, MER.**

I. Introduction

The deployment of digital television system in Brazil leads to modification of current transmission and reception standards, which implies the need for replacement of transmitters and antennas currently used by television broadcasters as well as the television sets installed in homes of television viewers [1].

The purpose of this study is to create and simulate an urban transmission environment to enable the analysis of the majors distortions suffered by the digital signal in the communication channel of the brazilian digital television system, ISDB-Tb. It was verified that there is not many studies at this topic [2]. The main metric used in this work is the Modulation Error Ratio (MER) measured at the receiver, which determines the relationship between the received symbol average power and its error average power in the received constellation. The MER measure observes the received symbol position at the demodulated constellation and the analysis of these values will determine the transmission channel quality [3]. The great majority of the measurement equipaments provide the MER and BER measurement (Bit Error Rate) separately, leaving aside the valuable information of the channel quality that the a joint analysis could bring to light.

The main causes of distortions in urban environments are the signal shadowing by natural or artificial obstacles, the Doppler effect, the path fading and the multiple interferences originated, mainly, at analog and digital transmissions systems with channels allocated at the same frequency or at adjacent ones [4] [5] [6].

II. The Proposed Simulation

The main simulated situations of this work consider the transmission channel of mobile and portable device content, since it makes no sense to evaluate fixed devices in movement. This channel will be called mobile channel or 1seg channel during this work. However, considerations about the transmission channel of contents for fixed set-top boxes, which will be called fixed channel or fullseg channel, won´t be neglected and are commonly found during this text. This is mainly due to the fact that the program of analysis of signals ISDB-Tb, installed in the spectrum analyzer, which displays the demodulated constellations, not display them in separate graphs. Thus, it became convenient to analyze the fullseg channel (64-QAM modulation) in this work. The parameters chosen were isolated and for each one was determined their influence on the degradation of the quality of the received constellation.

Given this scenario, the main variables of the mobile channel analysed at this work are:

- Received power;
- Speed of the mobile device;
- Propagation delay;
- Components of fading;
- The C/N Relation.

For each of these listed variables, an study of their relation with the modulation error ratio (MER) will the final result of this work.

Figure 1 shows off the complete setup of the measurement environment installed at Positivo Informática S/A digital TV laboratory. The equipments used were:

- A ISDB-Tb transmitter;

- A fading generator;
- A spectrum analyser;
- A mobile receiver (1seg) and a fixed receiver (fullseg);
- A power splitter.

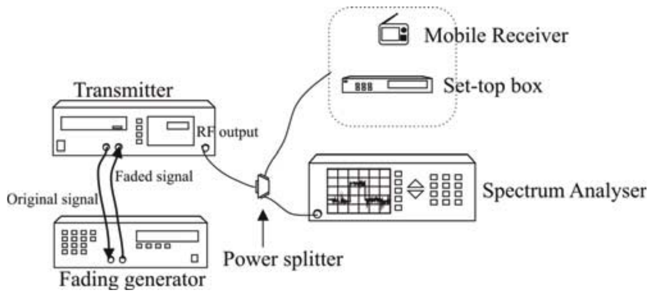


Fig. 1. Setup of the measurement environment

This setup illustrated at Figure 1 works on the following way:

- The transmissor generates the signal on an intermediate frequency and sends it to the fading generator;
- The fading generator, by its time, adds the particularly chosen distortions for each simulation case and sends it back to the transmissor at the same intermediate frequency.
- The transmitter generates the signal on an intermediate frequency and sends it to the fading generator; the spectrum analyser through a high quality coaxial cable passing through a power splitter.
- One of the power splitter outputs is connected to one the mobile terminal (sometimes to the fixed terminal) and the other is connected to the spectrum analyser, where the ISDB-T Demodulation Analysis software is installed.
- The mobile receiver is plugged in a notebook equiped with a video and audio software decoder. The streams contents were visualized at an proprietary application.

The ISDB-Tb Demodulation Analysis software has the functionalities of exhibition of the demodulated constellation, spectrogram and the measurement of the modulation error ratio.

III. The Modulation Error Ratio

The Modulation Error Ratio (MER) is the measurement of the degradation intensity of a modulated signal, which affects the receiver ability to recover the transmitted information. The MER can be similarly compared with the signal-to-noise relation on analog transmissions. This measure is very used on cable digital television systems due to its efficiency to express the combined effects of different perturbations at the communication channel. The MER reflects very well this combinations and is defined on a N symbol interval as follows [7],

$$\text{EVM} = \frac{\sqrt{\frac{1}{N}\sum_{j=1}[(I_j - \tilde{I}_j)^2 + (Q_j - \tilde{Q}_j)^2]}}{|\vec{v}_{max}|}, \qquad (1)$$

where I_j and Q_j are, respectively, the phase and quadrature components of the j-th received symbols and $\tilde{I}_j$ and $\tilde{Q}_j$ are, respectively, the ideally demodulated phase and quadrature components of the j-th received symbols. The calculation of the MER compares the current position of the received symbol and its ideal position. The value of MER increases when the symbols move away from its ideal position.

The combination of all interference in the transmission channel cause deviations in the position of the constellation symbols in relation to their nominal positions. Thus, this deviation can be considered as a parameter for measuring the magnitude of interference. And this is, in fact, the role of the modulation error ratio [8].

IV. Simulations

A. Case I: Received Power

In this simulation the signal to noise ratio used was 40 dB, which represents a fairly high level and practically nonexistent in practical situations and 40 dB, which represents a good reception environment. The purpose of using a so loud ratio was to isolate the behavior of the degradation of the modulation error ratio only depending on the reduction of the received signal power (P_R). Figures 2 and 3 shows demodulated constellallations. The simulation was done lowering the received power of -10 dBm to -90 dBm with a variation of -5 dBm for the first situation and -10 dBm for the second in each interval. This channel has no external disturbance.

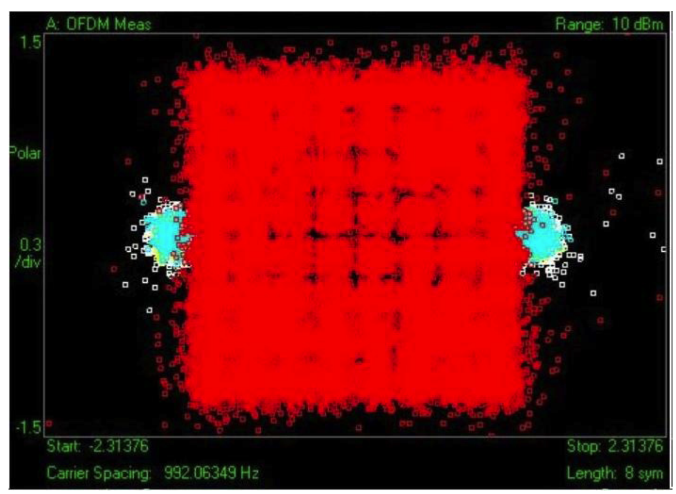


Fig. 2. P_R = -30 dBm and C/N = 20 dB

B. Case II: Multipath Fading Components

For this simulation case a transmission channel with a signal-to-noise ratio C/N of 40 dB and received power was -20 dBm was configured. From this, twenty multipaths were gradually added, one-by-one, at the fading generator. Figure 4 shows one sample of this simulation case.

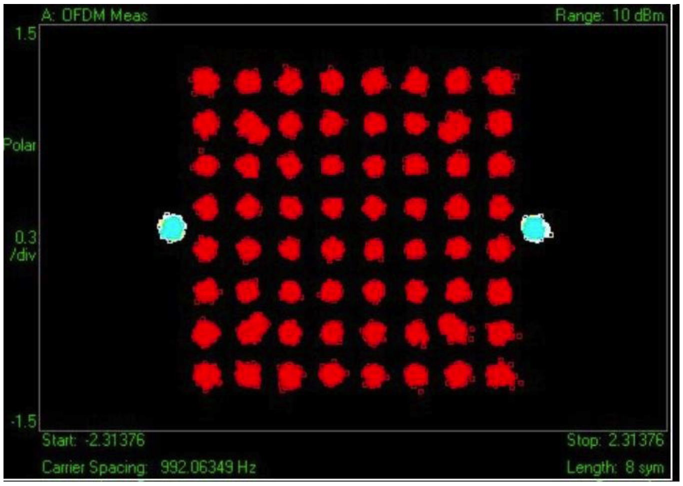


Fig. 3. P_R = -30 dBm and C/N = 40 dB

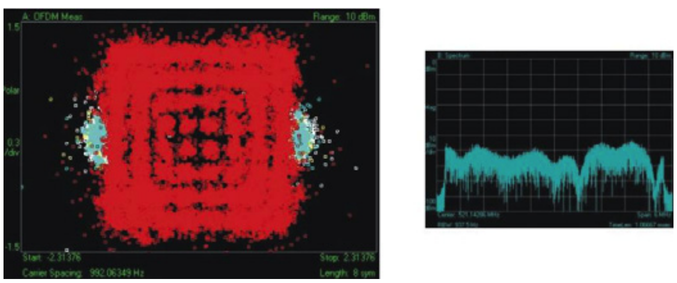


Fig. 4. Sample of the demodulated constellation of the channel with 7 fading components. C/N = 40 dB

C. Case III: Mobile Terminal Speed

For this simulation case a transmission channel with a signal-to-noise ratio C/N = 40 dB, received power was -25 dBm and five multipath fading components with significant power level was configured. From this, the speed of the mobile terminal was gradually increased at the fading generator. The Table IV-C shows a evaluation of the received video quality at a moving mobile terminal. Figure 5 shows a sample of the speed´s test.

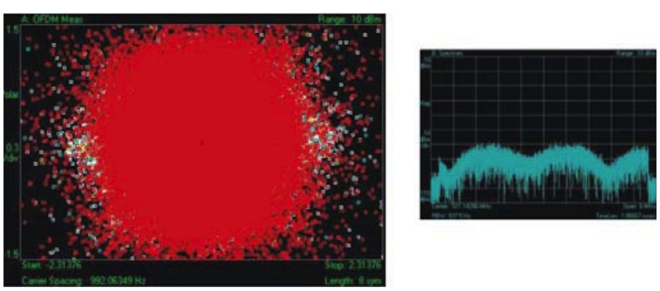


Fig. 5. Sample of the received constellation at 50 km/h and a screen capture of the spectrum. C/N = 40 dB

D. Case IV: Propagation Delay Spread

For this simulation were set two situations. In the first configuration, the received power was -40 dBm, which characterizes a common value of of the signal power found in practice, at good reception locations. In the second situation, the received power is was -80 dBm, which characterizes the futher places or bad condition of reception (e.g. strong multipath fading) for mobile terminals. This value is close to the limit of sensitivity of reception of the majority of mobile devices tested. Figure 6 shows the sample of the evolution of the degradation of the channel depending on the delay spread for each simulated situation.

Speed (km/h)	Mobile Channel	Fixed Channel
3	No issues	Many issues
20	No issues	Many issues
50	No issues	No signal
80	No issues	No signal
100	No issues	No signal
120	No issues	No signal
150	No issues	No signal
200	Few issues	No signal
300	Many issues	No signal
350	No signal	No signal

TABLE I
EVALUATION OF THE RECEIVED VIDEO AT A MOVING MOBILE TERMINAL

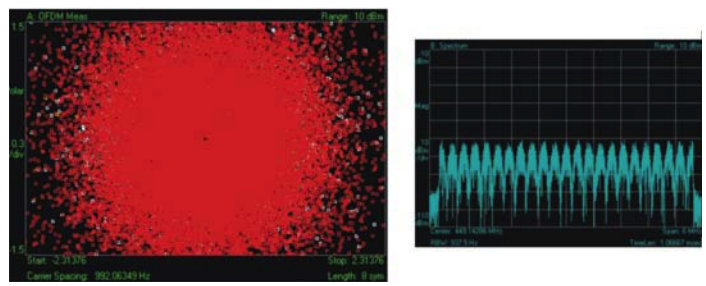


Fig. 6. Sample of the received constellation with a delay spread of 6 ms. P_R = -30 dBm and C/N = 40 dB

V. Results

A. Case I: Received Power

Figures 7 and 8 shows the resulting graphics of the analysis of the relationship between the received power and the modulation error ratio (MER). It is possible to see that for both simulated cases the MER has a nearly proportional degradation to the level of received power to -50 dBm. After this level, the degradation becomes more constant. It is worth mentioning that the Brazilian standard specifies that the threshold of sensitivity for receiving devices is fixed to -77 dBm. The Brazilian standard did not determined the level of sensitivity for mobile devices yet, but in laboratory tests with some devices, the threshold of reception for mobile devices is ranging from -85 dBm to -93 dBm for a signal-to-noise relation of 20 dB.

B. Case II: Fading Components

Figure 4 shows the graph of the relationship between the number of fading components and the MER. In this case, the

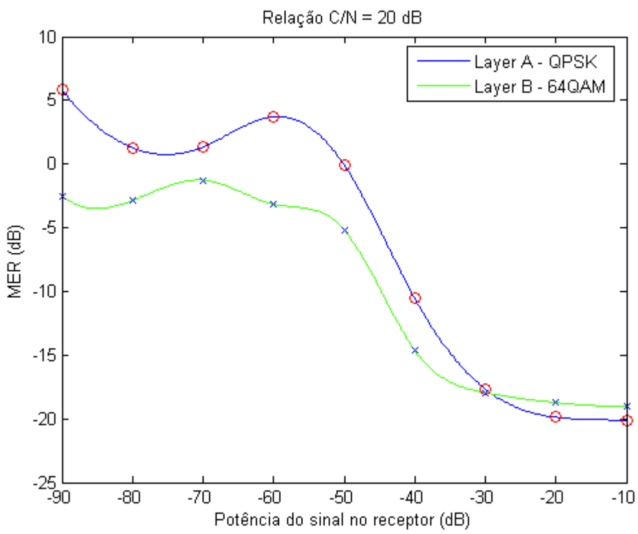


Fig. 7. Received Power × MER. C/N = 20 dB

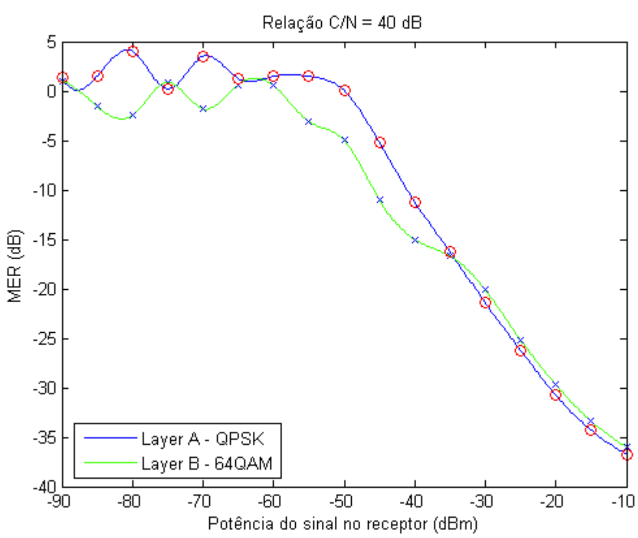


Fig. 8. Received Power × MER. C/N = 40 dB

samples showed a significant variation between the initial and final value, but showed a stable mean behavior in the range. Figure 4 depicts that even in a channel with a signal to noise ratio of 40 dB and a power of approximately -20 dBm, which represents a good condition for the reception, the influence of the quantity of fading components may cause the device not display the received content. This scenario is common when the transmitters are located in centers of large cities. In the field tests conducted in São Paulo it was noted that in several places at Paulista Ave, where the great majority of transmissors are located, even with a high level of received power, the quantity of fading components signal combined can cause the saturation of the receiver´s tuner. Another significant disturbance in this environment is adjacent channel interference from analog and digital transmission.

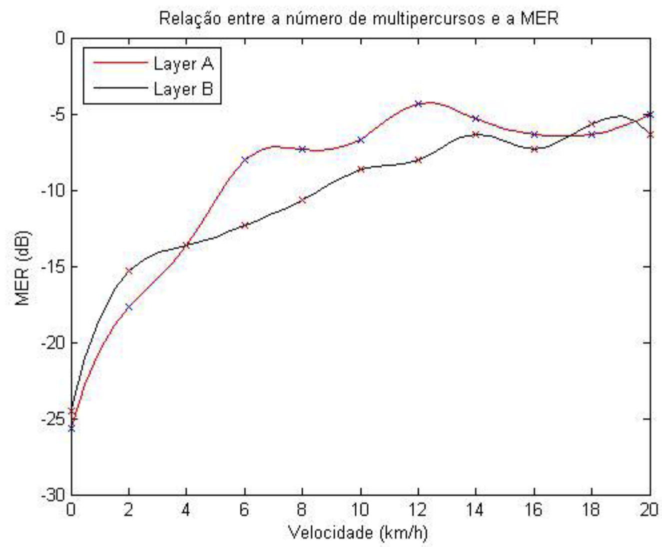


Fig. 9. Fading components quantity × MER. C/N = 40 dB

C. Case III: Mobile Terminal Speed

Figure 10 depicts the graph of the relationship between speed and MER. Also, Figure 10 indicates that the MER for the layer B tends to stabilize after 50 km/h. Anyway, according to the information in Table IV-C, at this speed would be difficult to demodulate the information from this channel since this layer has the purpose of transmission for fixed set-top boxes. In other words, it will be useless. The mobile channel (Layer A) indicates also a tendency to stabilize after the 100 km/h. Table IV-C, obtained in simulations in the laboratory with mobile devices, shows that a mobile device compatible with the Brazilian standard would jeopardized their reception at speeds above 200 km/h.

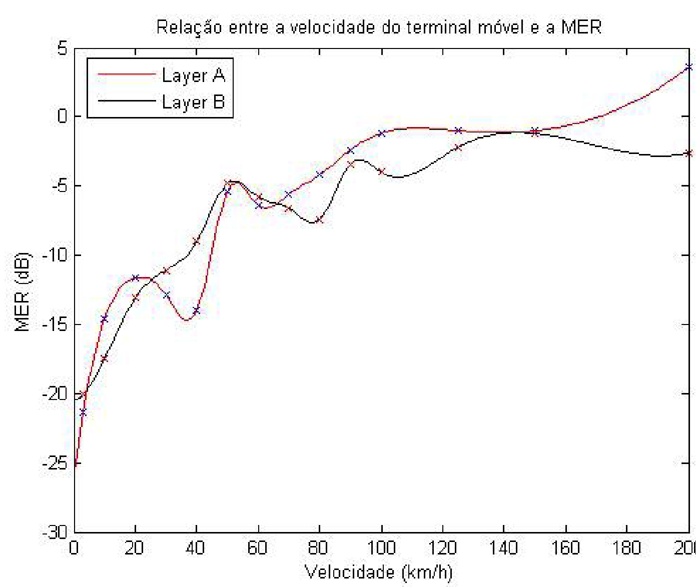


Fig. 10. Mobile terminal speed × MER. C/N = 40 dB

D. Case IV: Propagation Delay Spread

Figure 11 shows the graph of the relationship between the delay spread of a significant component of the signal and the MER. The duration of the delay interval used in the tests was from 1 μ to 6 μ. Through the curves of the graph, is possible to see that with a delay spread of, only, 6 μ the modulation error ratio decreases of approximately 5 dB. This is a considerable value in terms of lower received power, near the limit of sensitivity.

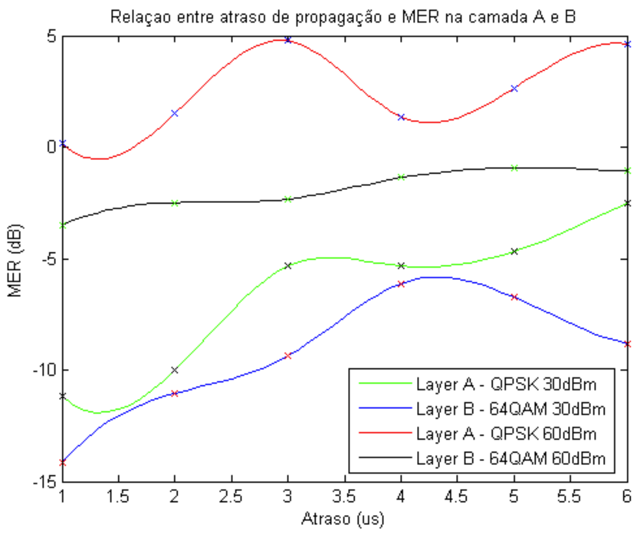


Fig. 11. Delay spread × MER. C/N = 40 dB

E. Case V: C/N Relation

Figure 12 shows the graph of the relationship between the modulation error ratio and the carrier to noise ratio C/N of the communication channel. In this case two situations were simulated with different transmission powers. In the first test used was -20 dBm as power transmission and the second test used was -40 dBm. Figure **??** indicates that to a carrier to noise ratio of approximately 12 dBm, the two simulated situations had a linear improvement in the modulation error ratio. From that point, the graph of Figure 12 shows that for the simulation case using -40 dBm, even increasing the value of the C/N, the value of the modulation error ratio MER tends to be constant. However, for the simulation case using -20 dBm, the increase of the modulation error ratio tends to vary almost linearly as a function of the improvement of the C/N relation.

VI. Conclusions

The graphs showed that, in practice, for the ISDB-Tb system, the QPSK modulation performance is most affected with the effects studied. However, the small number of symbols used in transmission and power, which implies on a greater distance between the symbols in the QPSK constellation comparing with the distance of 64-QAM constellation symbols. Thus, the QPSK modulation has a better immunity to the effects studied than the 64-QAM used for fixed reception

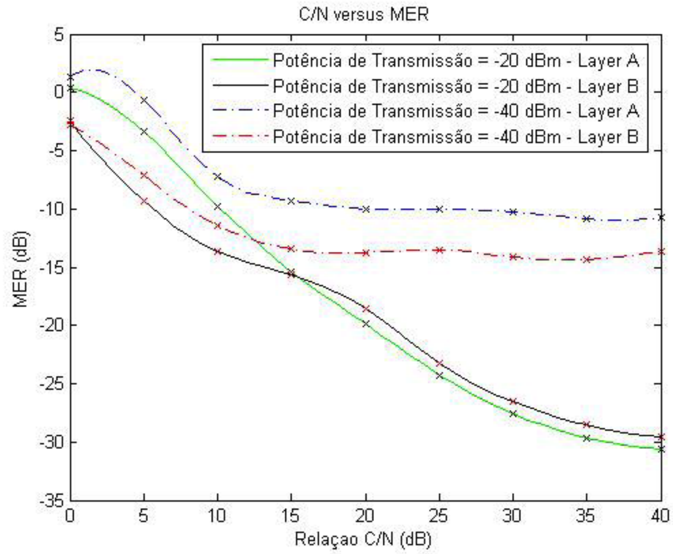


Fig. 12. C/N × MER

devices, even with elevated values of the modulation error ratio. Anyway, the study of the impact of the behavior of the variables over the modulation error ratio provides a better understanding of the degradation of the constellation for each case. It was observed that, even with high power and a high carrier to noise relation, the degradation of these variables implies, in most cases, the loss of the device ability to tune a digital channel.

However, to observe all the imperfections in the transmission channel, a joint analysis of the behavioral BER and MER is strongly recommended. One of the weaknesses of the MER is that its measure does not portray intermittent errors that result in an significant bit error rate.

References

[1] J. N. de Carvalho, "Propagação em áreas urbanas na faixa de UHF: aplicação ao planejamento de sistemas de TV digital," Master's thesis, Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica, Rio de Janeiro, Brasil, Agosto 2004.

[2] L. E. A. de Resende, "Desenvolvimento de uma ferramenta de Análise de desempenho para o padrão de TV Digital ISDB-T," Master's thesis, Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica, Rio de Janeiro, Brasil, Julho 2004.

[3] O. Mendoza, "Measurement of EVM (Error Vector Magnitude) for 3G Receivers," Master's thesis, International Master Program of Digital Communications Systems and Technology, Ericsson Microwave Systems AB, Mölndal, Sweden, Rio de Janeiro, Fevereiro 2002.

[4] F. S. C. A. L. T. R. Gunnar Bedicks Jr., Fujio Yamada and E. L. Horta, "Handheld Digital TV Performance Evaluation Method," *International Journal of Digital Multimedia Broadcasting*, vol. 45, no. 3, p. 5 páginas, Junho 2008.

[5] M. S. Alencar, *Televisão Digital*. São Paulo: Editora Érica, 2007.

[6] M.-S. A. Marvin K. Simon, *Digital Communication over Fading Channels:A Unified Approach to Performance Analysis*. John Wiley Sons, 2004.

[7] H. v. R. Walter Fischer, *Digital Television: A Practical Guide for Engineers*. Springer, 2004, 384 paginas.

[8] e. M. C. Y. Nasser, J.-F. Hélard, "System Level Evaluation of Innovative Coded MIMO-OFDM Systems for Broadcasting Digital TV," *International Journal of Digital Multimedia Broadcasting*, p. 12 páginas, Março 2008.

Decision Support for Monitoring the Status of Individuals

Fredrik Lantz, Dennis Andersson, Erland Jungert, Britta Levin
FOI (Swedish Defence Research Agency)
Box 1165, S-581 11 Linköping, Sweden
{flantz, dennis.andersson, jungert, britta.levin}foi.se

Abstract: *Systems for monitoring status of individuals are useful in many situations and for various reasons. In particular, monitoring of physiological status is important when individuals are engaged in operations where the work load is heavy, e.g. for military personnel or responders to crises and emergencies. Such systems support commanders in the management of operations by supporting their assessment of Actors' physiological status. Augmentation of the commanders' situation awareness is of particular importance. For these reasons, an information system that supports monitoring of such operations is presented. The system gathers data from multiple media sources and includes methods for acquiring data from sensors, for data fusion and decision making. The system can also be used for after action review and training of actors.*

1.INTRODUCTION

For a large number of reasons, it is important to monitor the physiological status of individuals subjected to high physical workload in situations that may lead to exhaustion and reduced performance. Such situations concern soldiers in military operations, fire fighters and other responders to different crises and emergencies that face high workload situations. However, novel methods and technologies must be developed to make the system effective and efficient. Examples of such technologies comprise development of a wireless body area network (WBAN) for the individual, i.e. body worn sensors and equipment for wireless communication. Of crucial importance to all such systems is that they should be easy to carry around by the individuals, as well as efficient with respect to how data are collected, analyzed and transmitted to the end-users for further analysis in their decision making processes. Furthermore, means for integration of data from multiple data sources must also be available. This requires further development of techniques and methods for sensor data analysis, multi-sensor data fusion and techniques for search and selection of relevant information from the data sources. In all, the collected information should be used as input to various decision support tools that must be developed to support the users in their monitoring, planning and decision making activities. Eventually, physiological monitoring systems must also be possible to use in conjunction with various command and control (C2) systems.

The structure of this work is outlined as follows. The objectives of the work are presented in section 2. Section 3 presents and discusses the fundamentals of the work, which includes the physiological aspects and the general system focus. Communication issues are discussed in section 4 while the means for data integration, i.e. data fusion, are discussed in section 5. In section 6 are the architecture of the command and control system, the decision support tools and the system for after action review presented. Related works are discussed in section 7 and finally the conclusions of the work appear in section 8.

2.OBJECTIVES

In this work, a system for handling multimedia information for physical monitoring of individuals is presented. Two aspects are in main focus for this work. The first aspect concerns the methods and the algorithms for collection and analysis of physiological information for determination of the status of the actors. The system must support the decision makers' situation awareness by collection, fusion, filtering, and visualization of data adapted to the users requests. The second aspect concerns the development of a system architecture for such a monitoring system. Another aspect of interest is to support after action review (AAR) [10] to give the actors feedback from training sessions or actual missions. However, the actual development of the WBAN is not within the scope of this work.

3. FUNDAMENTAL REQUIREMENTS

3.1 Physiological aspects

Physiological and psycho physiological monitoring can be of interest for various types of applications such as health and safety monitoring, medical emergencies, physically challenging exercises, and study of task performance.

Continuous supervision of human physiological status requires a set of sensors capable of detecting the variables of interest. Depending on the target application these variables may differ significantly. Health and safety monitoring usually focus on observations of one or more critical factors, such as the heart activity in a patient with diagnosed heart failure or the potentially fatal heat stress for a fire fighter. For a medical emergency it is important to use sensors capable of detecting vital signs such as body temperature, respiration, heart rate, and blood pressure. In a physiologically strenuous situation in a hostile environment it may be relevant and feasible to measure for instance body and ambient temperature, heart rate, perspiration, altitude, position, and body posture. Determination of task performance often comprises both physiological measures of fitness as well as psycho physiological measures including subjective ratings, heart rate and heart rate variability indicating mental stress.

3.2 System aspects

In order to assess physiological status the various variables need to be properly recorded and further processed. The data recording system must be designed to minimize interference with the users' activities and their ability to move around freely. Long duration exercises and difficult environments put additional and tough requirements on the sensors and the recording system. Generally, sensors should be durable and easy to apply while the recording system must be built to assure low weight and volume, flexibility, and low power consumption. The overall system structure is described in Figure 1. Data are transmitted wirelessly from the actors to the personal server where data are processed and further transmitted to the decision support system where the information is further processed and visualized.

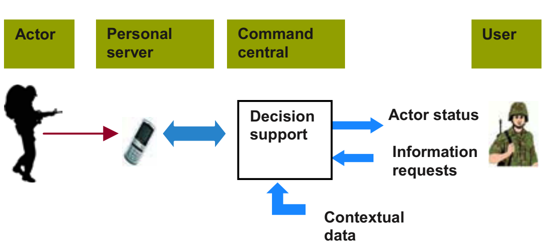


Figure 1. The overall structure of the system.

4. COMMUNICATION

Decision makers can be located in a command central that may be located at a significant distance from the monitored individuals. This implies that communication between the personal server and the command central must be executed via existing communication infrastructures. For the system in this study, all communication is performed via Internet by attaching a GPRS module to the personal server.

GPRS communication is relatively expensive in terms of energy consumption compared to computations in the personal server. Since one of the design goals of this system is low power consumption there is a need to minimize the amount of data being transferred. There are several ways of reducing GPRS communication as discussed below.

4.1 Data reduction through fusion

By fusing the data at an early state the data being transferred can significantly be reduced. This implies that computations should be done locally on the personal server and that the variables of interest are known. An example of this is if the system automatically determines the body posture of the actor rather than transmitting e.g. raw accelerometer data. The downside of such a solution is that it limits the possibilities for post action analysis.

4.2 Data reduction through skipping

Data may be collected at high sample rates, sometimes much higher than needed for the analysis, and skipping samples may be an option. When two consecutive samples have no significant difference then there is no need to transfer the second sample. What is considered as an insignificant difference is application dependent and may be changed at runtime by the user if need arises. Skipping could also be executed on a regular basis by sending every *ith* sample which will reduce the granularity in the data collected at the command central.

4.3 Data reduction through subscription

In a scenario where several individuals are being monitored and/or many sensors are being used on each individual it is unlikely that all data are needed at all times. Analysts may have different needs for different stages of the operations. A subscription solution would then help reducing the data flow since the analysts always can subscribe to only those data they are currently interested in. Thus, sensor data not subscribed to will not be transferred.

5.DATA FUSION

Data fusion is the process for combination of data to estimate or predict the state of some system. The process is commonly separated into five different functional levels; Sub-object assessment, object assessment, situation assessment, impact assessment and process refinement, see [8]. The end-product of the data fusion process is a situation picture that is common for all the sensors and other data sources that has been used to estimate the state.

5.1 Modeling of physiological processes

According to [4], the models that are used in modeling physiological phenomena are often linear, deterministic and non-dynamic in spite of the fact that these phenomena often are non-linear, stochastic and dynamic. Consequently, there is large room for improvement in this area using common techniques from the data fusion area, e.g. Dynamic Probabilistic Networks, Hidden Markov Models or Sequential Monte-Carlo methods. As most models are aimed at a certain group of people (e.g. females of a certain age and weight), it is also possible to improve the effectiveness of the systems by tailoring the algorithms/systems to the unique characteristics of the individual actors, see also [7].

5.2 Data fusion for actor state estimation

The state of the actors is a joint description of several status variables of interest in the particular application. Position, velocity, motion mode (i.e. running, standing, lying down, etc.) and heart rate are fundamental variables of interest in many applications. These variables must in some cases be determined through the combination of data from several sensors. For instance, by combining data from accelerometers and GPS motion mode can be determined. In the data fusion process, the uncertainties in the data are taken into account. Data are weighted according to their certainty and erroneous data can be identified and excluded. Other variables can be included in the actor state as described in chapter 3.1.

5.3 Aggregated measures of actor status

An aggregated measure of actor status should be an indicator of the ability of the actors to perform their tasks. Different measures, using different sensors and data fusion methods, therefore need to be used depending on the application and the actors' tasks. In some applications the amount of work performed by the actors are an effective measure of the status. In e.g. [11], a Physiological Strain Index is calculated based on the heart rate and the core temperature. It is important to note that the determination of the status of a healthy individual can be more difficult than for an injured/sick individual since the healthy individuals' status values can be expected to be less extreme than for injured/sick individuals. In the current situation, it is also important to take the actors and their co-workers own evaluation of their status into account. An automatically deduced status value may in many applications only serve to notify the user of a situation where he/she must request information from the actors.

5.4 Context and data fusion

The context where the actors are performing their tasks is important for interpretation of the status values. For instance, it is important to know the motion mode and velocity of the actors in order to interpret other physiological values correctly, e.g. their pulse. The geographical context, the weather conditions as well as the equipment carried and clothing worn by the actors are also crucial to the interpretation of the status values. In the data fusion process, these data must be collected and fused with the actors' state values. An example is the usage of the 3D terrain models that can be used to improve the estimation of the altitude of the actors. Conversely, the actors' state can be used to interpret the context, e.g. if it can be detected through the motion pattern of the actors that a certain area is difficult to traverse, it may consequently be classified as difficult.

5.5 Automatic alarms

One of the most important functions of a decision support system for monitoring the status of individuals are functions to relieve the users from having to continuously monitor sensor data. An important component in such a system is therefore algorithms for automatic detection of the actors or the group states that deviate from the normal or expected, i.e. *anomalies*. Using algorithms for anomaly detection, the users can be left with the task to verify alarms given by the automatic algorithms and take required actions when appropriate.

Development of adequate algorithms for anomaly detection is very much a research issue. Normal values are, for instance, heavily dependent on the context and on the task performed by the actors. In some cases, what is "normal" can be defined by the status of other actors, while in other situations an individual model must be used. Consequently, the system must also allow for individual variations.

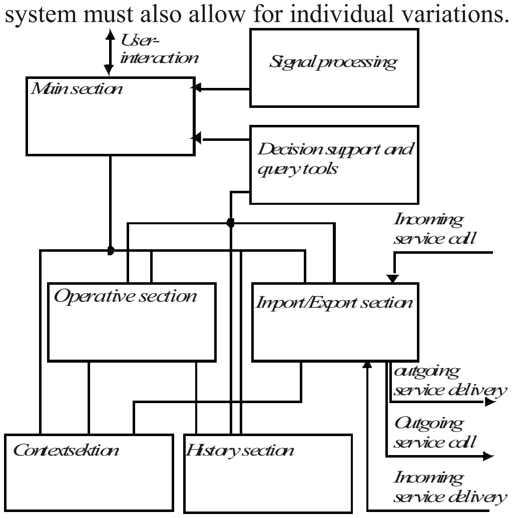


Figure 2. An overview of the system architecture.

6. SYSTEM ARCHITECTURE

6.1 Command and control architecture

The C2 systems architecture discussed here is service oriented and consequently highly modular. In particular, the modular approach taken has its roots in the approach originally taken in the work demonstrated by Jungert and Hallberg [12] and the variation here is adapted to the monitoring of the status of individuals. The system is based on what is called the *role concept model* in which the basic concepts are: (1) *views*, (2) *services*, (3) *roles*, and (4) their relationships. Primarily, the model is developed to provide for mission support in command and control processes. The model illustrates how users relate to their role in the information management process. Views are made up by services and *visuals,* where a visual corresponds to a visualization of a view instance. The role concept model is further discussed in [5] and [12].

The basic building blocks of the architecture are the *sections*. The most important sections and their relations can be seen in Figure 2. To each section a number of views are assigned corresponding to various specialized services and supported by one or more visuals. Sections can be replaced when required. The two most important sections are the operative section and the main section. The main section, which does not contain any views, contains the main interface of the system through which the users can manipulate the views and their content by means of the available services attached to each of the occurring views. The operative section includes a set of views that are of vital importance to the ongoing work as they include the currently available operative information; thus the views in the operative section represent the current operational picture.

The Import/Export section is basically a buffer for incoming and outgoing information, which due to given service calls, made either by a local user or by an external user, can be sent or received. The incoming information generally contains sensor data from groups of individuals being monitored by the system. The context view (CXV) in the context section is the storage point of all available background information such as maps. The user can, by means of the context view, define the required area of interest (AOI) and display it in the current operative view in the operative section resulting in what here is called a view instance. Eventually, the view instance of the current operative view (COV) is completed through the overlay of either the individuals of interest or the groups of interest to the current mission. A view instance can then successively be updated resulting in new instances.

The history section hosts the history view (HYV), which can be seen as a repository for all view instances created prior to the current operative view instance presently residing in the current operative view.

The current operative section, which is the most important and powerful section contains four views for support of the operative work in the monitoring process; these four views are:

- Current operative view (COV)
- Physiological information view (PIV)
- Individuals of interest view (IIV)
- Groups of interest view (GIV)

COV, which displays the view instance corresponding to the current operative picture, can be directly interacted upon. For instance, by clicking at the icons of the individuals in the view instance, physiological information corresponding to any group or single individual can be made available in PIV. To allow for more complicated results this may also be combined with a query language, see further below. IIV and GIV show individual and group information respectively for personnel subject to monitoring. This information may include physiological and as well as location information.

Most available services in the system are part of the views, of which some are simple and in many ways similar to ordinary systems commands while others correspond to conventional services. Three main groups of services have been identified: 1) *view handling services*, e.g. create a new view, 2) *view* *instance handling services, e.g import or create a new view instance* and 3) *view instance manipulation services, e.g. update the content of an existing view instance*. The type is depending on the target of the service, see the given examples below. The number of needed services is fairly large and some are unique to a certain view while others occur in more than one view. Because of the large number of services, only some limited examples can be given here to illustrate the service concept. An example of a view instance handling service is:

- Request specified information (a view instance) from a user and store the information in IMV.

Another service for view instance handling is:

- Go back and display the view instance of COV created at time t.

This view instance is accessed from HYV and displayed in COV.

An example of a view instance manipulation service is:

- Update COV with information from IMV

The decision support section of the systems architecture still requires further research efforts. However, it will eventually also contain some type of query tool. In earlier work, a query language for heterogeneous sensor data, called ΣQL, see e.g. [6], was developed. This query tool also has capabilities for sensor data fusion. To be used in this environment ΣQL needs to be modified and simplified. This is mainly due to e.g. group leaders of the individuals subject to monitoring who for this purpose are using PDAs to present the monitored information. Thus, the objective here is to adapt the query interface to a query structure related to dynamic queries [1] but also to make it suitable for PDAs as described and demonstrated in [3]. The purpose of the query tool is generally to use information available in COV, IIV and GIV as input to the queries and produce the requested output, either as tables or as graphs, in PIV.

As a consequence of the service based approach , which allows import and export of information from all participating users the content of the four views of the operative section together corresponds to a shared operational picture [13], which forms the basis for the decision making process.

6.2 After action review

AAR [10] is a formalized method for evaluation of exercises and operations. The F-REX method and tolls [2] support this type of procedure through introduction of Reconstruction & Exploration [17].

F-REX supports after-action reviews by enabling visualization of any data type in a chronological order and related to other data sources, giving the analysts the opportunity to quickly get an understanding of the data being observed in relationship to the context in which it was sampled. Figure 3 shows such a context from a rescue services exercise in Sweden 2006 [2]. The system visualizes and plays back concurrent data from several data sources dispatched over a large area enabling the users to quickly get an overview of the current situation at different locations and thus get a relevant context for the analyses.

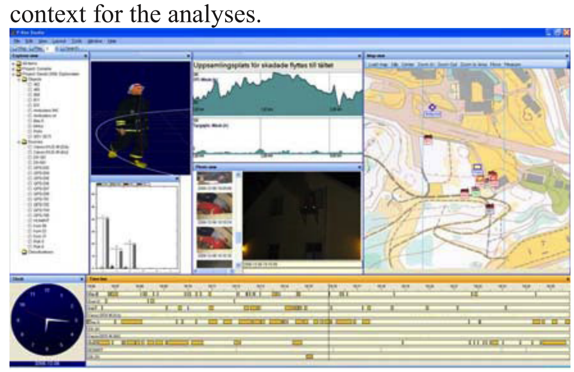


Figure 3. Synchronized visualization of sampled data and multimedia information for contextual analysis in F-REX Studio. The current layout displays a priori information, timeline, photos, video, a map with GPS tracks, heart rate, altitude, stance and statistical metrics.

F-REX partly implements the system architecture described above. Figure 3 shows several visuals synchronized automatically to a COV and with contextual information provided by the CXV (i.e. the map in the GPS track visual, statistics for the chart visual and photo/video in the multimedia visual). The timeline at the bottom of the screenshot provides an interface to easily access the HYV while IIV and GIV can be setup using the tree structured a priori interface to the left. By selecting individuals or groups of individuals the heart rate, visual, map and stance visuals will be updated to reflect the current selection.

This tool is ideal for post action analysis of data being gathered by the monitoring system. Further, extensions to allow online visualization of data in F-REX are being planned to allow the system being used as a real time decision support tool and not just in post real time reviews.

7. RELATED WORK

The research literature exhibits a large number of works where the monitoring of the individuals' physiological status is in focus. Generally, this type of monitoring is also used in many different applications. However, related work of particular interest to the work discussed here concerns primarily systems and systems architectures where physiological monitoring of individual and groups of individuals is of concern. Integration of such systems to command and control systems is another important issue to deal with. Other literature of interest concerns decision support tools used in this context.

McGrath et al. [16] discuss a crisis management system called ARTEMIS with the primary purpose to improve the care of wounded soldiers. The system is a part of a command and control system and it has also been developed with the intension to improve the users' situation awareness, which is carried out by improved information gathering even under severe situations. The authors also argue that, from this perspective, more reliable decisions can be taken.

AID-N [9] is a triage system with command and control capacity based on SOA (service oriented architecture). AID-N is thus a service oriented approach. The system exploits shared data models to support data exchange between different heterogeneous subsystems. The system can be seen as a test bed for improved co-operation between crisis management organizations. A powerful aspect of the system is that it through its service architecture allows for a simplified distribution (sharing) of data between users of the subsystems.

Among the different monitoring systems the work by Lin et al. describes a system called RTWPMS [14], which is a mobile system supporting examination of patients where physiological information is measured by means of sensors; e.g. for measuring of blood pressure and temperature. Another example of a monitoring system is described by Lorincz et al. [15], [19]. This system corresponds to a surveillance system using a sensor data network for data gathering. The primary applications of concern fall in the areas of crisis management and medical surveillance of patients. In this system, simple queries can be put as well. Another monitoring system that relates to the work discussed here is described in [22].

An example of a system for extensive medical decision making and which also uses methods for sensor data fusion is discussed in [20]. Staceyand et al. [21] describes a system that can perform intelligent analysis on clinical data. In [18] a network approach for measuring physiological parameters is discussed.

8. CONCLUSIONS

In this work, a system for monitoring the physiological status of individuals has been discussed. Physiological parameters are measured by means of sensors attached to the bodies of the individuals in focus for the monitoring process. The parameters are then transferred and further analyzed in a system for determination of the individuals status; this system will eventually also be integrated with a command and control system that can be used for both military and civilian applications. Of importance is also the integration of an after action review system. The purpose of the latter system is to offer techniques and methods to give the users improved means to judge the consequences of certain operations in which the individuals are involved, but also to see how these individuals react to the given circumstances. In the decision support component this may be combined with geographic information, weather and other relevant information.

Other aspects that will need further attention in future research will be concerned with the development of methods for automatic alarms through anomaly detection. Methods for tracing the general state of the individuals in combination with their motion patterns are also of concern.

REFERENCES

1. Ahlberg, C., Williamson, C., Shneiderman, B., *Dynamic queries for information exploration: an implementation and evaluation*, Proceedings of the Conference on Human Factors in Computing Systems (CHI 92), ACM Press, New York, 1992, pp. 619–626.
2. Andersson, D., Pilemalm, S. & Hallberg, N., *Evaluation of crisis management operations using Reconstruction and Exploration*, Proceedings of the 5th International ISCRAM Conference, May 4-7, 2008, Washington, DC, USA.
3. Burigat, S., Chittaro, L.,,*Interactive visual analysis of geographic data on mobile devices based on dynamic queries*, Journal of Visual Language and Computing, Vol. 19, No 1, February, 2008, pp 99-122.
4. Carson, E., Cobelli, C., *Modeling Methodology for Physiology and Medicine*, Academic Press San Diego, CA, USA, 2001.
5. Chang, S.-K., Jungert, E., *A Self-Organizing Approach to Mission Initialization and Control in Emergency Management*, Proceedings of the

International Conference on Distributed Multimedia Systems, San Fransisco, CA, September 6-8, 2007.

6. Chang, S.-K., Jungert, E., and Li, X., *A Progressive Query Language and Interactive Reasoner for Information Fusion*, J. of Information Fusion, Elsevier, Vol 8, no 1, 2006, pp 70-83
7. Committee on Metabolic Monitoring for Military Field Applications, *Monitoring Metabolic Status: Predicting Decrements in Physiological and Cognitive Performance*, National Academic Press, Washington, DC, USA, 2004.
8. Hall, D. L., Llinas, J., (Eds.), *Handbook of multisensor data fusion*, CRC Press, New York, 2001.
9. Hauenstein, L., Gao, T., Sze, T. W., Crawford, D., Alm, A. and White, D., *A cross Service-oriented Architecture to Support Real-Time Information exchange in Emergency Medical Response*, Eng. in Medicine and Biology Soc. (EMBS '06), 28th Annual Intern. Conf. of the IEEE, New York, NY, Aug. 2006.
10. Headquarters Department of the Army, *A Leader's Guide to After-Action Reviews* (TC 25-20), Washington, DC, 30 September 1993.
11. Hoyt, R., W., Buller, M., Zdonik, S., Kearns, C., Freund, B., Obusek, J., F., *Physio-Med Web: Real Time Monitoring of Physiological Strain Index (PSI) of Soldiers During an Urban Training Operation*, RTO HFM Symposium on “Blowing Hot and Cold: protecting Against Climatic Extremes”, Dresden, Germany, 8-10 October, 2001.
12. Jungert, E., Hallberg, N., *An Architecture for an Operational Picture System for Crisis Management*, Proceedings of the 14th Inter.l Conf. on Distributed Multimedia systems, Boston, MA, Sept. 4-6, 2008.
13. Jungert, E., *A Theory on Management of Shared Operational Pictures for Command and Control Systems Design*, IADIS Int. Conf. on Information System 2009, Febr. 25 - 27, Barcelona, Spain.
14. Lin, B.-S., Chou, N.-K., Chong, F.-C., Chen, S.-J., RTWPMS: *A Real-Time Wireless Physiological Monitoring System*, IEEE Trans. on Information Techn. in Biomedicine, Vol. 10, No 4, Oct. 2006, pp 647-656.
15. Lorincz, K., Malan, D. J., Fulford-Jones, T. R. F., Nawoj, A., Clavel, A., Shnayder, V., Mainland G. and Welsh, M., Moulton, S., *Sensor Networks for Emergency Response: Challenges and opportunities*, IEEE Computer, Oct.-Dec. 2004, pp 16-23.
16. McGrath, S. P., Grigg, E., Wendelken, S., Blike, G., De Rosa, M., Fiske A. and Gray, R., *ARTEMIS: A vision for Remote Triage and Emergency Management Information Integration*, Dartmouth University; Nov.. 2003.
17. Morin, M. *Multimedia Representation of Distributed Tactical Operations*, Linköping Studies in Science and Technology, Dissertation No. 771, Linköping University, Linköping, Sweden, 2002.
18. Rahman, F., Kumar, A., Nagendra, G., Gupta, G. S., *Network Approach for Physiological Parameters Measurement*, IEEE Trans. on Instrumentation and Measurement, Vol 54, No 1, Febr. 2005, pp 337-346.
19. Shnayder, V., Chen, B.-R., Lorincz, K., Fulford-Jones, T. R. F. and Welsh, M., *Sensor Networks for Medical Care*, Technical Report TR-08-05, Div. of Engin. and Applied Science, Harvard University, 2005.
20. Sintchenko, V., Coira, E. W., *Which clinical decisions benefit from automation? A task complexity approach*, Int. J. of Med. Informatics, vol. 70, 2003, pp 309-316.
21. Staceyand, M., Mcgregor, C., *Temporal abstraction in intelligent clinical data analysis: A survey*, J. of AI in Medicine, Vol. 39, No 1, Jan. 2007, pp 1-24.
22. Yu S.-N. and Cheng, J.-C., *A Wireless Physiological Signal Monitoring System with Integrated Bluetooth and WiFi Technologies*, Proceedings of the 27th annual IEEE Conf. on Eng. in Medicine and Biology, Shanghai, China, Sept. 1-4, 2005, pp 2203-2206.

Assessment of IT Security in Emergency Management Information Systems

Johan Bengtsson, Jonas Hallberg, Thomas Sundmark, and Niklas Hallberg

Abstract—**During emergency management the security of information is crucial for the performance of adequate and necessary operations. Emergency management personnel have commonly only novice skills and interest in IT security. During incidents they are totally preoccupied with the crisis management. Hence, the security mechanisms have to be well integrated into the emergency management information systems (EMIS).**

The objective of this paper is to illustrate how security assessment methods can be used to support decisions affecting the information security of EMIS. The eXtended Method for Assessment of System Security (XMASS) and the accompanying Security AssessmeNT Application (SANTA) are introduced. The method and tool support the security assessment of networked information systems capturing the effects of system entities as well as the system structure. An example is provided to illustrate the use of the method and tool as well as the importance of effective firewalls in networked information systems.

Index Terms— **IT security, IT security assessment, Emergency management.**

I. Introduction

When emergencies occur, there is little time to consider other issues than how to handle the situation at hand. Focus is required in order to minimize the negative consequences of the situation. Critical decisions have to be made based, often, on uncertain information. Thus, the decisions have a significant impact on the success to handle situations. The information used as foundation for the decisions are more commonly generated, communicated, processed, provided and interpreted by the use of information technology (IT) based systems, i.e., emergency management information systems (EMIS). EMIS are decision support systems to be used in all parts of emergency management and response [1]. They support the emergency managers in planning, training and coordinating operations [2]. EMIS can be used to, e.g. display and analyze possible event locations, available resources, transportation routes, and population at risk [3]. EMIS have the potential to dramatically increase our ability to, foresee, avert, prepare for and respond to extreme events [4]. However, with extensive use the dependency in EMIS increases and, consequently, the need for trusted and reliable EMIS. Thereby, IT security issues are vital to consider for the information systems support to be used for emergency management. Thus, it is essential to have a valid understanding of the security posture of EMIS. A serious problem is posed by the fact that if there is no method to establish the current level of security in EMIS, then there is no way to decide whether the IT security levels of these systems are adequate. Furthermore, the effect of any actions to improve the IT security will be unknown.

Thus, it is crucial to design methods that will remove the ad hoc nature of security assessment for EMIS. In this paper, a structured method for the assessment of EMIS is presented. The method has been implemented as a tool, which is used to assess security levels of coalition networks at the Combined Endeavor, an international communications and information system interoperability exercise.

M.Sc. Johan Bengtsson, johan.bengtsson@foi.se
Dr. Jonas Hallberg, jonas.hallberg@foi.se)
M.Sc. Thomas Sundmark, thomas.sundmark@foi.se
Dr. Niklas Hallberg, niklas.hallberg@foi.se

All with the Swedish Defence Research Agency, Linköping, Sweden

II. Background

This section presents IT security, IT security assessment and the context of the study.

A. IT security

IT security, also referred to as computer security, is defined in many different ways depending on the context. Excellent descriptions of various aspects of IT security are provided by, e.g., Anderson [5], Bishop [6] and Gollmann [7]. Consequently, it is hard to give an explicit definition, which is suitable for all contexts. Gollmann [7] states that there are several possible definitions, such as, "deals with the prevention and detection of unauthorized actions by users of a computer system." In this paper, the term IT security relates to upholding the characteristics of confidentiality, integrity, and availability of IT systems and the data processed, transmitted, and stored in these systems.

B. IT security assessment

Assessment of IT security is performed in order to establish how well IT systems meet specified security criteria, based on measurements of security relevant system characteristics or effects. Hubbard [8] points out that in order to measure something; it has to be distinctly clear what it is that should be measured. However, measurements do not have to yield exact results. Successful measurements improve the knowledge about the studied phenomena [8]. Hence, IT security assessments are to provide knowledge about the security of IT systems. This knowledge can be used to support, e.g.:

- the comprehension of the current security posture by the actors responsible for the IT security,
- the development and operation of information systems, e.g. EMIS, with adequate security levels,
- risk management,
- training and awareness concerning IT security,
- the communication of IT security issues,
- security management, and
- trust in IT systems [9].

Although IT security deals with technical elements, comprehensive IT security assessments need to consider other related aspects, such as the organizational, human, and contextual aspects. The inclusion of these aspects emphasizes the need to consider their influence on the security levels of systems. However, IT security assessments do not include the assessment of the security of organizations, persons, and contexts themselves.

Several approaches to security assessment have been presented. Security metrics programs refer to the process of:

- identifying measurable system characteristics and effects,
- measuring these security characteristics and effects, and
- produce illustrative, comprehensive presentations of the results [10-12].

Adequate security metrics should be consistently measured, inexpensive to collect, expressed by numbers, and have a unit, such as seconds [10]. The interpretation of specific security metrics is left to the user. Proponents of security metrics programs claim that the characteristic of triggering discussions on the meaning of the presented results is a key benefit. In contrast, the approach presented in this paper, the eXtended Method for Assessment of System Security (XMASS), aims at providing system-wide security assessment values including the effects of system structure and inter-connections [13,14]. Thus, the whole system is considered during the assessment rather than isolated system characteristics or effects.

Attack-based methods assess systems based on the steps that adversaries have to complete in order to achieve their goals, e.g., [15,16]. The method based on the weakest-adversary security metric aims to enable the comparison of different system configurations based on the attributes required to breach their security [16]. Characteristics of network configurations and the current attack stages, e.g. root-level shell access on a specific host, form the states of the system models. The transition rules describe the requirements for and consequences of the transitions from system states into other system states. Describing the actual prerequisites of successful attacks, the presented results are intuitive. However, the analysis of results may not be as straightforward, e.g., when making comparisons of the system effects resulting from different system configurations. The XMASS does not require the knowledge of specific vulnerabilities that can be used to penetrate systems; instead assessments are based on the security qualities of systems.

Methods based on system characteristics combine values of selected characteristics to produce security values which represent the security levels of complete systems. The Security Measurement (SM) framework is used to estimate scalar security values [17]. In order to transform relevant security characteristics into measurable system effects or characteristics a decomposition method is described. The outcome is a tree with measurable security characteristics as leaves. For the aggregation of security values, the weights and mathematical functions capturing the relations between the nodes in the resulting tree have to be decided. Because of the generality of the method, large efforts are required to design specific methods based on the framework. Since assessments based on the XMASS can utilize different sets of security characteristics to capture the security levels of systems, the process of systems modeling is more clearly specified. Like security metrics programs, the SM framework lacks support for capturing the security effects of system structure, which is explicitly supported by the XMASS.

C. *Study context*

The Combined Endeavor constitutes an extensive communications and information system interoperability exercise. The participants are members of the North Atlantic Treaty Organization (NATO) and the Partnership for Peace (PfP). During Combined Endeavor 2007 Sweden participated with equipment in, and connected external networks to, the established Region B network. This network is used as the target of evaluation in this paper.

III. METHOD FOR ASSESSMENT OF IT SYSTEM SECURITY

In order to assess the security of systems, it is essential to capture the underlying characteristics and effects related to the systems as well as defining how the computation of security values should be performed. Thus, both the systems to be assessed and the computations to be performed have to be modeled. Provided these models, the base data has to be captured and the aggregated values have to be computed in order to receive the final assessment results. To benefit from the produced results, their presentation has to be adapted to the recipient (Figure 1).

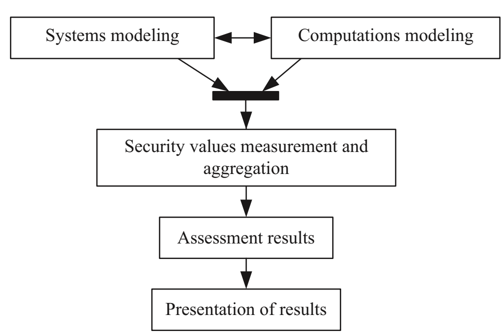


Figure 1: The outline of methods for security assessment.

The eXtended Method for Assessment of System Security (XMASS) [13,14] has been formulated according to the structure presented in Figure 1 and to fulfill the following requirements.

- Provide users with relevant data on the IT security posture of networked information systems.
- The effects of system entities as well as the system structure should be captured.
- Since there is no fixed definition of IT security, the method should support the assessment of different security aspects which together conjure the definition of IT security of the user.
- The method should be flexible in order to support the diverse needs of different users.
- The reuse of assessment data should be supported.

In XMASS, assessments are based on the available knowledge regarding the security characteristics of the system entities and their relations [13]. The system modeling is supported by the possibility to create profiles for standardized system entities and their relations. There are no explicit limitations in the method regarding which system entities can be modeled.

The computation of higher-level security values is controlled by the computations model, which can be specified by the users, but is tied to the structure of the system. Thus, the computation of aggregated security values, not just the input, depends on the system models as well as the computations models. The assessment results are presented for individual entities, for entities in a system context, and for the entire system.

A. *XMASS tool*

The tool implementation of XMASS is based on the NTE (New Tool Environment) [18], which is a software framework supporting the implementation of security assessment methods. NTE supports the definition of Requirement Collections (RCs), which enable the specification of different security features. These security features can in turn be broken down into a number of security requirements. NTE simplifies the implementation of tools for security assessment methods by providing basic functionality such as:

- file handling for organizing systems and projects,
- a data access layer to provide a simple way of reading and writing to the database, and
- well defined interfaces to facilitate the implementation.

The actual systems modelling and assessment functionality is implemented as a plug-in for the NTE, called SANTA (Security AssessmeNT Application). The SANTA is designed to facilitate variation of values and settings, which makes it possible to evaluate the XMASS and improve its functionality. An example of this is that the security-related values of a modelled entity are structured as a profile which can be reused by other entities of the same type. A change in one profile, affects all entities using that specific profile.

B. *System security assessment workflow*

System security assessments in the SANTA are performed according to the workflow illustrated in Figure 2. A white background indicates that the activity is part of the calculation modelling, while a blue background indicates that the activity is part of the system modelling. The workflow consists of five activities: (1) Create Requirement Collection, (2) Create templates, (3) Create profiles, (4) Create system model and (5) Perform system assessments. The activities are described in the following sections.

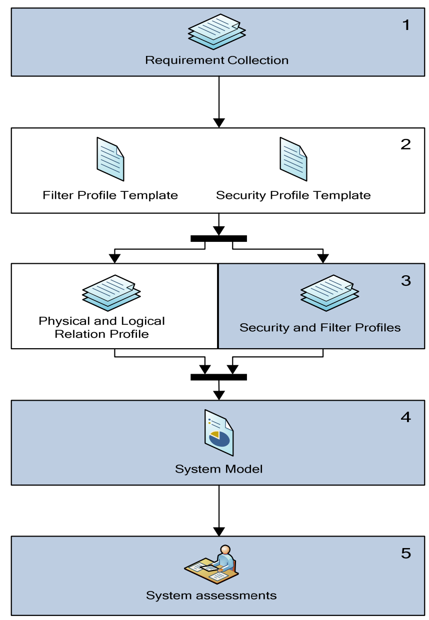


Figure 2: The workflow for security assessments.

1) *Create Requirement Collection*

A Requirement Collection (RC) is a specification of the security features that should be regarded during the security assessment. Each security feature is mapped to a set of security requirements. The fulfilment of these security requirements will decide the security values of systems and system entities corresponding to this security feature. Higher security values for a security feature indicate that the feature is adequately supported by the assessed system or system entity. The RC is the basis for the security assessment since it specifies what needs to be fulfilled in order to receive favourable assessment results. The templates and profiles created in the following steps are all dependent of the specified RC.

2) *Create templates*

The security profile template defines the importance of each requirement specified in the RC. The requirements of each security feature are divided into two categories; fundamental requirements and important requirements. Fundamental requirements have to be fulfilled in order for the assessed entity to be considered as fulfilling the security feature. The important requirements, on the other hand, are prioritized regarding their relative importance. The prioritization is performed with a method based on the criteria weighting used in the Analytic Hierarchy Process, AHP, [19] and decides to what extent each requirement affects the security value of the regarded security feature. It is possible to regulate the maximum total influence of the important requirements.

The filter profile template defines how the specified network traffic filtering functionalities affect the security value of each security feature. The relative influence of the filtering functionalities is, for each security feature, specified with the help of the method based on the AHP [19]. It is possible to specify the maximum effect a traffic filter can have on each security value.

3) Create profiles

A profile is a grouping of values which concerns one or more entities or relations. The main reason for grouping values into profiles is to facilitate the modelling and simplify the variation of values, i.e., an alteration of a profile affects all entities or relations using that specific profile. There are four types of profiles: (1) security profiles, (2) filter profiles, (3) physical relation profiles, and (4) logical relation profiles.

There are two main types of entities defined in the XMASS, traffic generators and traffic mediators. A traffic generator is an entity which generates traffic and can for example be a workstation computer or a server. A traffic mediator is on the other hand an entity which only mediates traffic and can for example be a router or a switch. Each entity in a system has a security profile which describes to what degree the entity fulfils the security requirements specified in the RC. A fulfilment value of 1 indicates complete fulfilment of a requirement, while 0 indicates non-compliance. A fulfilment value between 0 and 1 indicates partial fulfilment.

The filtering functionality and capability of different traffic mediators can differ widely. Therefore filter profiles are used to specify how the filtering of the mediator affects the system security.

Relations are described using relation profiles. There are two types of relation profiles; one for physical relations and one for logical relations. The physical relation profile differs from the other profiles by being specified as a system-wide setting. Hence all physical relations in a system are modelled using the same physical relation profile. The physical relation profile describes associations between entities interconnected through physical means such as wired or wireless communication. The logical relation profiles are, on the other hand, specified per relation and describe logical relations such as VPN tunnels etcetera.

4) Create system model

Once the previous three steps have been completed, it is possible to start with the visual modelling of the system. Entities and relations are created by simply clicking, dragging and dropping in the modelling surface. When creating a new entity the first step is to choose whether to create a traffic generator or a traffic mediator. For the traffic generator, only a security profile needs to be selected, while the traffic mediator needs a filter profile as well. When creating a relation, either the physical relation profile or a suiting logical relation profile is selected. To support the modelling of extensive systems, it is possible to specify sub-systems that can be instantiated in the visual system model.

During the visual modelling of the system, it is essential that all the necessary profiles are available. If any profiles are missing, the third step of the process has to be revisited.

5) Perform system assessments

Once the computation modelling and the system modelling have been completed the system assessment can start. The foundation for the security values produced by the XMASS is the System-dependent Security Profiles (SSPs) that are computed for all the traffic generators in the system model. The computation of the SSPs depends on the specified computation and system models.

The SANTA offers different ways to extract assessment results from the system model. Next to the modelling surface is a panel showing the calculated system-dependent security values which are aggregated security values reflecting the system as a whole (Figure 3).

For more advanced assessments of a system, there is a built-in evaluation tool which makes it possible to generate graphs of how changes of security values affect the security. This can for example be used to illustrate how the security values are affected if the filtering policies of used firewalls are changed.

There is also a built-in tool for calculating how much each entity affects the security of each other entity in the system. This tool can for example be used to identify weak spots in the system, i.e. the entities having the worst influence on the other entities in the system.

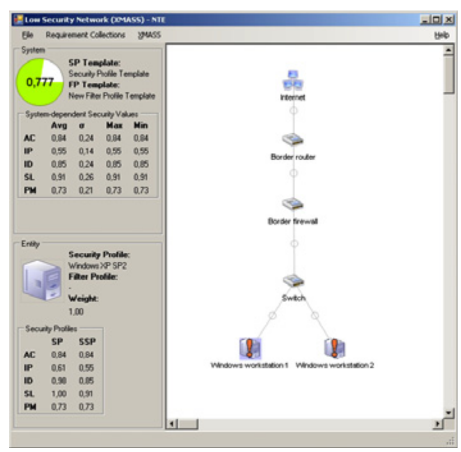


Figure 3: An overview of the SANTA.

IV. Modeling of the CE07 network

The modeled system used in this security assessment is, as mentioned earlier, Region B of the network used at the Combined Endeavor 2007 (CE07). The graphical view of the SANTA model of the network is presented in Figure 4.

The purpose of the designed network was to connect the subnets of the participating nations to a core network in order to allow them to communicate with each other and the other nations connected to the core network. The participating nations controlled their own subnets, so while designing the Region B network the security focus was set on the firewalls in between the subnets. All firewalls used in the network were of the model Färist, which is used by the Swedish Armed Forces. For a specification of the hardware used in the Region B network along with the requirement collection, templates, profiles and settings used in the model refer to [20].

There are three different types of symbols used in the model representing *traffic mediators*, *traffic generators* and *subnets of traffic generators*. A subnet represents a given quantity of identical traffic generators interconnected through a switch. In this network each subnet represents ten workstations using Microsoft Windows XP SP2. Information about the actual number of workstations per subnet was not available at the time of the modelling.

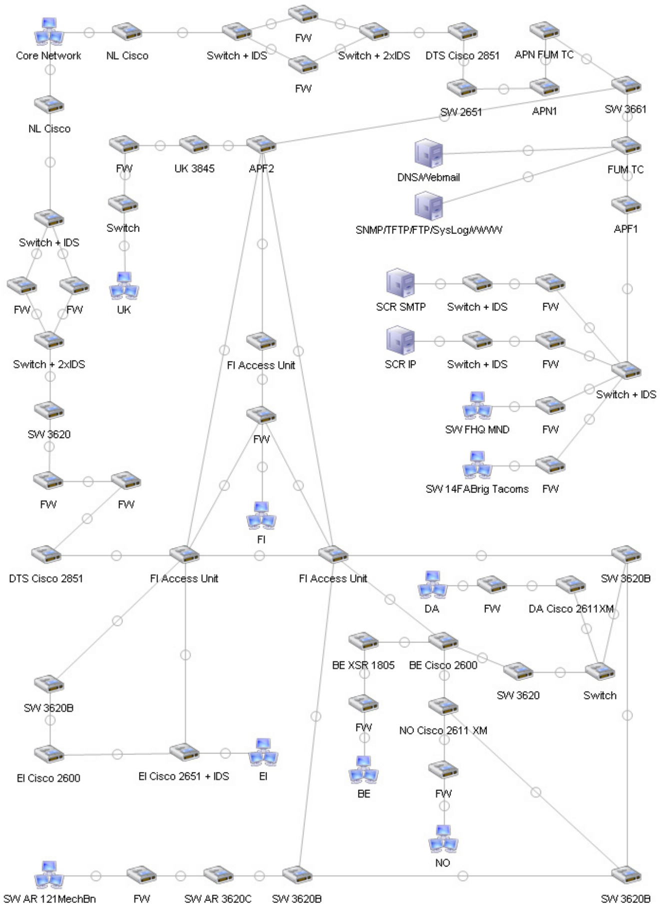


Figure 4: The model of the CE07 Region B network.

V. Security assessment of the CE07 network

As mentioned earlier, the firewalls are central for the security level of the CE07 network. To illustrate the importance of the firewalls, a security assessment is performed for different levels of traffic filtering. The requirement collection used for the security assessment is the collection of requirements on security mechanisms used by the Swedish Armed Forces [21]. This collection regards the security features *Access Control*, *Intrusion Prevention*, *Intrusion Detection*, *Security Logging* and *Protection against Malware*. In Figure 5, the graph represents the system-wide security profile, i.e., an aggregation of all the entity SSPs in the system model. The security values are plotted in the graph where the filtering capabilities of the firewalls are linearly increased from zero to the maximum level of the firewall.

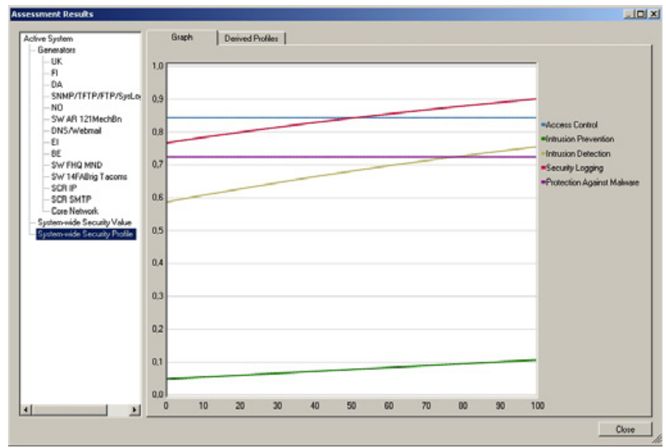


Figure 5: Assessment results.

To further illustrate the importance of the firewalls, an incident occurs when an unprotected wireless router is connected to the network. This makes the network open for unknown, and probably also unwanted, clients having an unknown level of security. This threat has been modeled as a subnet of ten traffic generators having the lowest possible security level connected to the network at the same switch as the UK subnet (Figure 6).

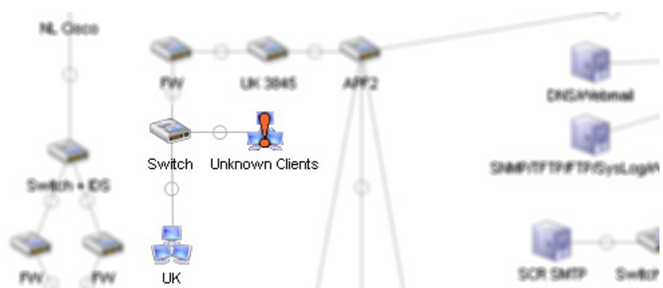


Figure 6: Changes made to the network.

By performing the same security assessment, as with the original model, the importance of the firewalls is even more obvious (Figure 7).

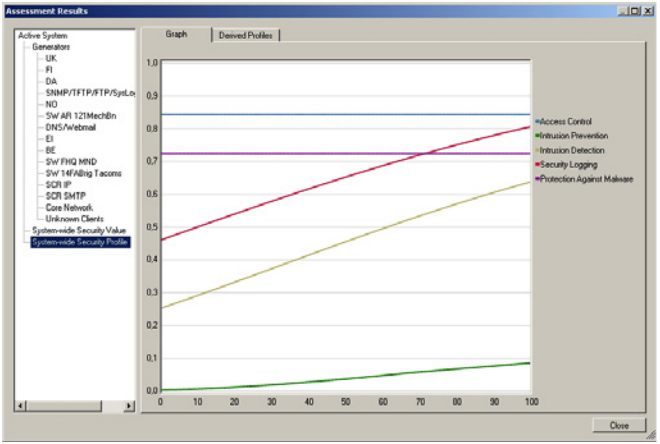


Figure 7: Assessment results for the modified network.

VI. Discussions

In emergency management many critical decisions are based on information obtained by the use of information systems [2,4]. To obtain adequate and effective emergency responses, it is crucial that emergency managers can trust and rely on provided information. Hence, the ability to ensure a sufficient level of IT security within emergency management information systems (EMIS) is essential. This can be achieved through methods and tool for IT security assessments. This paper presents the method XMASS and the tool SANTA enabling the assessment of IT security.

In XMASS, the assessments capture the effects of system entities as well as the system structure. The CE07 example presented in this paper illustrates how filtering affects the security levels in large networks. As can be seen from the results presented in Figure 5 and Figure 7, the security values for AC and PM are constant. This is because these security features have been modeled to be independent of the security level of the other system entities. The security values for IP are generally low when no filtering is active in the firewalls. This is because the IP value of the security profiles is relatively low and there are in total many entities collectively affecting the values of the SSPs. When the filtering capabilities of the firewalls increase, the security values corresponding to the SL, ID, and IP improve. This is because the non-perfect values of the neighbors shielded of by firewalls increase due to filtering.

The importance of filtering is illustrated by the fact that the relative difference between the security values of the networks, with and without the unknown clients connected through the unprotected wireless router, decreases with more effective filtering. For example, considering the ID security feature, the security value decreases with 58% when there is no filtering and 16% when the maximum filtering of the modeled firewalls is assumed.

The usability of EMIS as support for the decision making, within emergency management, requires the integrity as well as availability of critical information. Modern EMIS are network-based and connected to public networks. Hence, the traffic filtering capability of EMIS is one crucial aspect in order to reach and maintain both integrity and availability. This aspect is regarded in the assessments performed with XMASS and SANTA. Hence, such methods and tools support the design, configuration and operation of trustworthy and reliable information systems for emergency management.

VII. References

[1] M. Kwan and J. Lee, “Emergency Response after 9/11: the potential of real-time 3D GIS for quick emergency response in micro-spatial environments,” *Computers, Environment and Urban Systems*, vol. 29, 2005, pp. 93-113.

[2] D. Ozceylan and E. Coskun, “Defining Critical Success Factors for National Emergency Management Model and Supporting the Model with Information Systems,” *Proc. 5th International Conference on Information Systems for Crisis Response and Management ISCRAM2008*, F. Fiedrich and B. Van de Walle, eds., Washington, DC, USA: 2008, pp. 276-83.

[3] S. Pilemalm and N. Hallberg, “Exploring Service-Oriented C2 Support for Emergency Response for Local Communities,” *Proc. 5th International Conference on Information Systems for Crisis Response and Management ISCRAM2008*, F. Fiedrich and B. Van de Walle, eds., Washington, DC, USA: 2008.

[4] J.R. Harrald, “Agility and Discipline: Critical Success Factors for Disaster Response,” *The ANNALS of the American Academy of Political and Social Science*, 2006.

[5] R. Anderson, *Security engineering: A guide to building dependable distributed systems*, Wiley, 2001.

[6] M. Bishop, *Computer Security - Art and Science*, Addison-Wesley, 2003.

[7] D. Gollmann, *Computer security*, Chichester: Wiley, 2006.

[8] D.W. Hubbard, *How to measure anything: finding the value of "intangibles" in business*, Hoboken, N.J.: John Wiley & Sons, 2007.

[9] N. Hallberg, J. Hallberg, and A. Hunstad, “Rationale for and Capabilities of IT Security Assessment,” *Proc. IEEE SMC Information Assurance and Security Workshop IAW '07*, 2007, pp. 159-66.

[10] A. Jaquith, *Security metrics: replacing fear, uncertainty, and doubt*, Addison-Wesley, 2007.

[11] D. Herrmann, *Complete guide to security and privacy metrics: measuring regulatory compliance, operational resilience, and ROI*, Auerbach Publications, 2007.

[12] E. Chew, M. Swanson, K. Stine, N. Bartol, A. Brown, and W. Robinson, *Performance Measurement Guide for Information Security*, National Institute of Standards and Technology, 2008.

[13] J. Hallberg, N. Hallberg, and A. Hunstad, *Crossroads and XMASS: Framework and Method for System IT Security Assessment*, Swedish Defence Research Agency, FOI, 2006.

[14] J. Hallberg, J. Bengtsson, and R. Andersson, *Refinement and realization of security assessment methods*, Swedish Defence Research Agency, FOI, 2007.

[15] B. Laing, M. Lloyd, and A. Mayer, “Operational Security Risk Metrics: Definitions, Calculations, and Visualizations,” *Metricon 2.0*, Boston: 2007.

[16] J. Pamula, S. Jajodia, P. Ammann, and V. Swarup, “A weakest-adversary security metric for network configuration security analysis,” *Proc. 2nd ACM Workshop on Quality of Protection*, 2006.

[17] C. Wang and W. Wulf, “A Framework for Security Measurement,” *Proc. National Information Systems Security Conference*, 1997, pp. 522-533.

[18] J. Bengtsson and P. Brinck, “Design and Implementation of an Environment to Support Development of Methods for Security Assessment,” Linköping University, Department of Electrical Engineering, 2008.

[19] T. Saaty, *Fundamentals of Decision Making and Priority Theory - with the Analytic Hierarchy Process*, Pittsburgh: RWS Publications, 1994.

[20] T. Sundmark, “Improvement and Scenario-based Evaluation of the eXtended Method for Assessment of System Security,” Linköping University, Department of Electrical Engineering, 2008.

[21] Swedish Armed Forces, *Requirements on security mechanisms (In Swedish: Krav på SäkerhetsFunktioner)*, Headquarters, 2004.

Practical Experiences in Using Heterogeneous Wireless Networks for Emergency Response Services

Miguel A. Sanchis, Juan A. Martínez, Pedro M. Ruiz, Antonio F. Gomez-Skarmeta
Dept. of Information and Communications Engineering
University of Murcia
{miguelsanchis, juanantonio, pedrom, skarmeta}@um.es

Francisco Rojo
112 Emergency Services of
Murcia Region
fco.rojo@carm.es

Abstract

Emergency services are one of the main application areas that can benefit from the adoption of wireless communication technologies. Unlike other application domains where all devices have very similar capabilities, we foresee that for emergency services the key is being able to build a heterogeneous wireless communication platform which encompasses the advantages of each of the different existing technologies. In this paper, we analyze the main requirements, describe a possible supporting architecture and report some real experiences in collaboration with our local emergency response service (112).

I. INTRODUCTION AND MOTIVATION

Wireless communications are currently being present in many applications in our daily life, as well as in many products and services. It is certainly one of the areas whose importance in the society is experimenting an unprecedented growth. On other hand, emergency response is also an area that is rapidly changing due to the continuous incorporation of new cutting edge communications technologies. So, it becomes natural to apply recent advances in wireless networking to the highly important area of emergency response and management. In particular, we are very interested in exploiting the easy deployment, increased bandwidth and huge reliability of recent technologies.

In our proposal, we consider the use of Wireless Mesh Networks (WMNs). They have emerged as a key technology for next-generation wireless networking, due to their easy deployment, ability to work with little infrastructure and survivability. WMNs are dynamically self-organized and self-configured, where nodes are automatically create an ad hoc network and maintain the mesh connectivity. We also consider the integration of WMNs with existing high-capacity wireless technologies such as WiMAX and UMTS. Our final goal is to be able to provide real-time images from the place where the emergency is taking place to the remote control center so that more informed decisions for the emergency management can be made.

In this paper, we analyze the requirements and solutions for efficient video distribution from first responders to the control center of the emergency service, in the particular case of our local emergency coordination centre (112, which is similar to 911 in USA). We also set up a demonstration prototype (Fig 1) and report some performance results to assess whether a heterogeneous wireless network can fulfill the needs of our local emergency services.

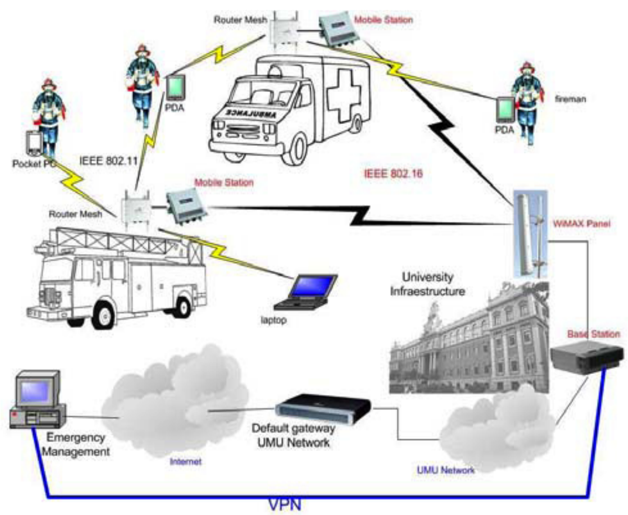


Figure 1. Logic structure of proposal architecture.

The rest of the paper is organized as follows. The first section describes some related work. Section II analyzes and describes the requirements to solve existing shortcomings in traditional emergency technologies. Section III addresses the main performance requirements which must be satisfied by the proposed communication architecture. The remaining three sections describe the proposed solution, the test-bed setup and lessons learnt and finally the experimental results from our experiments. Section VII concludes this paper.

II. RELATED WORK

In [1] several problems which appeared during the natural disaster happened in US because of Katrina Hurricane are analyzed in relation to communication, management and infrastructure supply. Main conclusions state that incompatibility of available technologies is identified as a great problem, and the lack of frequency bands, caused a low communication rate among responsible entities of emergency groups. These three deficiencies are identified as top priority issues which should be addressed in the design of communications architecture for emergency services.

There are some proposals for emergency situations based in WMNS. In [2] proposed a Hybrid wireless mesh network with particular application for emergency and crisis scenarios, where performance measurements are obtained during a real mesh testbed. In [3] authors present the GeoBIPS architecture (which using underlying wireless mesh networking technology) an overview of the different hardware and software components. The improvements carried out in this work have been the channel selection algorithm and the On The Go Coverage Indicator.

Based on the architectures proposed in the literature [2-5], and considering the general and our specific requirements based on the experience provided by the 112 emergency service of Region of Murcia, we concluded that a solution based on three layers is suitable to provide a wireless infrastructure connected to the wired backbone network linked with the control center. In addition the architecture is designed on the basis of different existing technologies, deployed for giving a wireless network with a maximum coverage and relying on specific emergency services.

III. PERFORMANCE REQUIREMENTS

As first priority, a public safety communication system must be able to provide reliable voice, image and video data to allow an acceptable rendering quality. The proposal presented in this paper comprises communication areas provided with a great bandwidth, a wider coverage and with an extended flexibility to adapt to deployment and topological changes.

The delay of messages transmissions from the source to the target devices has to be minimal, in order to deal with real time demands. In this kind of scenarios, this delay cannot be higher than 150 ms and not only for voice over IP but also for video streaming.

The communication bandwidth related to paths without broken links, has to be enough to offer video-streaming services. The recommended bandwidth is around 512 Kbps for combined audio/video streaming, taking into account that this value is only for one connection. Regarding the bandwidth required for video only transmissions, it depends on some factors such as image quality or frame rate. Our experience shows that the minimum video traffic rate is 384 Kbps for one video channel and 128 Kbps for two audio channels, when the image size is 320x240 pixels.

With the lens that video and sounds emission are not negatively affected, the latency obtained during this process, defined as the time difference between the moment in which a link fails and the moment in which an alternative router starts being used [6], has to be minimized for this kind of applications, since the quality of service could be suffer a negative impact.

Communication systems for dealing with emergency management must be reliable and robust, in order to operate under adverse situations and hostile environments. Robustness is just one of the main features of wireless mesh networks [7]. The communication platform must also be self-healing in order to react quickly to topology changes, since critical emergencies (like fires) require fast adaptive networks.

IV. PROPOSED ARCHITECTURE

The proposed architecture is based on a wireless mesh network at the road edge, and technologies are involved are WiMAX (802.16), and Wi-Fi (802.11a/b/g), to create the real mesh. It also considers 3G technologies for first response and establishment of the initial communication between terminals at the emergency place and the control center. The real tests carried out with the system demonstrate the advantages of this mesh network.

The scenario considered is composed by three important parts. First, the WiMAX deployment performs gateway functions for the nodes of the mesh network; Second, WiMAX clients are located at emergency vehicles and are directly connected to base station, which is placed in a strategic place to give a major coverage. These WiMAX clients are connected through the wired interface to mesh routers, which also brings wireless coverage to mobile devices used by emergency staff. The third and last element is the emergency center, where applications used for emergency management are hosted.

In the diagram of Fig. 1, there are two wireless technologies which operate at once. On one hand WiMAX is used to communicate each vehicle (trucks, cars, motorcycles) with the base station and the wired network. The WiMAX deployment is connected with the control center, using VPN, which provides a secure communication environment. On the other hand, a wireless router into vehicles is used to create a mesh, with multiple radio interfaces to exploit the available bandwidth. Since they use multiple radio interfaces, the simultaneous channels maximize the WMN capacity [8].

The logical structure presented in this document also includes 3G/2G cellular technologies, due to the rapid response they offer and their great expansion. Thus, the first response to an emergency situation could be notified from a mobile device, establishing an initial video connection with the emergency services center. As a response to this first action, the necessary resources are sent to the emergency scene (emergency vehicles with a WiMAX client). The incident area network is then established to offer Wi-Fi coverage to mobile devices, with a personal area network (mobile phones, laptops, PDAs, etc…). In this moment it is already possible to broadcast video in streaming and establish a common data connection between end devices and the service center. This way, the incident scope is accurately determined, and subsequent actions can be performed depending on the emergency.

WiMAX has been chosen since this technology represents an optimal solution for IP-based broadband wireless communications. Its capacities in terms of coverage, offered data rates and terminal mobility, makes WiMAX be easily applied on disaster scenarios. Furthermore, WiMAX fulfill QoS constraints in the field of emergency management scenarios [9]. On other hand, WMNs can be easily implemented through IEEE 802.11a/j-like radio interfaces, offering broadband capacity in emergency scenarios. An extra factor which determines the success of WMN technologies is interoperability, which is another major problem, identified in emergency response operations.

IEEE 802.16 and 802.11 must be considered both oriented to reach a full coverage of the disaster area, moreover to offer to users different connection alternatives. In that sense, in [10] authors proposed two interconnecting techniques between a Wi-Fi and a WiMAX network having into account the need to support the same QoS level between two networks.

In the real tests carried out we have worked closely with the emergency management services of Region de Murcia, conducting several tests for checking the correct operation of the video streaming transmissions. The purpose of this collaboration is to provide the necessary communication infrastructure, by means of WiMAX, to interconnect mesh network with the service central. The platform was found useful to manage incidences by means of live audio/video of a great quality, suitable to the first-person-lived incidents from the place of the disaster or emergency, managing to carry out the gesture of the available resources in a more efficient way.

V. TESTBED SETUP AND LESSONS LEARNT

As described above, the scenario is composed by elements of two different technologies, WiMAX devices and several mesh routers. In order to achieve our experiments we deploy our scenario using Alvarion BreezeAccess 4900 as the WiMAX devices and Gateworks Avila GW2348-4 as the mesh routers. The WiMAX client and mesh routers get the power supply by means of PoE described in the IEEE 802.3af standard.

The main problems we have faced related to mesh device are the followings: Firstly, the configuration of the Ad-hoc mode in the wireless interfaces, not only in the mesh routers but also in the mobile devices (laptops), since they were not very often easy to associated automatically. Secondly, it was the developing and installation of the routing algorithms into the Mesh routers. Due to the necessities of low energy consumption, they have different processor architecture (ARM). So the way to insert the algorithm in these devices is using cross-compiling technique. Finally, other obstacles were caused by the lack of information of some protocols and the configuration requirements for the kernel on these devices.

Regarding the WiMAX technology, the devices used in our testbed are not compliant with the mobile standard making difficult to get good marks in our scenario.

VI. EXPERIMENTAL RESULTS

This section shows some performance results obtained in the wireless mesh network and in the WiMAX link. Three performance tests have been carried out to verify the usefulness of the infrastructure over a real environment and also to check the efficiency of the routing protocol OLSR [11] integrated in wireless mesh routers. In order to get our experimental results we have used a CBR generator so as to get the maximum bandwidth according to our scenarios requirements. These scenarios are detailed in the following paragraph.

The first performance test was carried out using only the WiMAX technology. The target of this test was to obtain the maximum bandwidth recreating the situation explained above in the proposed architecture section, i.e., a real time audio and video broadcasting which shows the location of an incident. The second test with WiMAX and routers mesh is about to connect the client WiMAX with the router mesh in such a way that the first one is used as a gateway for the router mesh. The last experiment introduced one hop extra to the previous test using a new mesh router. Therefore, the communication from a laptop to the WiMAX client is transmitted along two mesh routers instead of using only one.

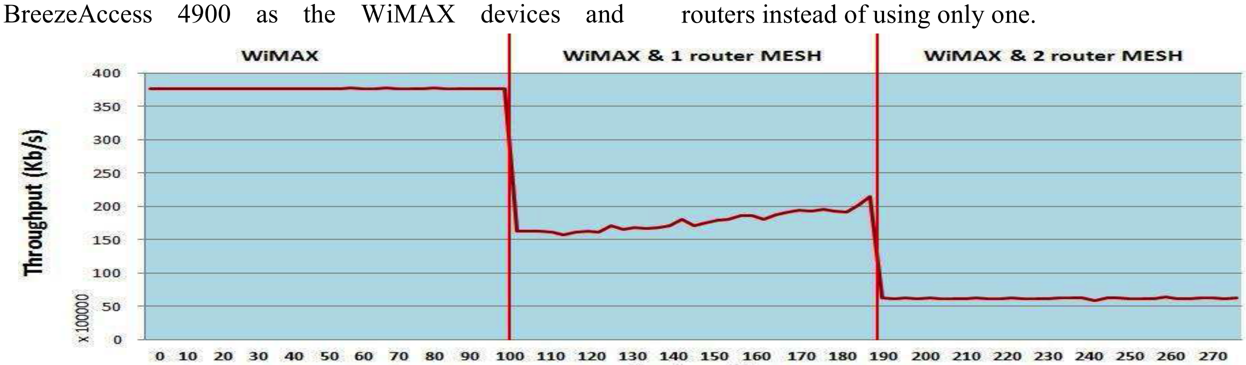


Figure 2. Average Throughput (Kbits per second)

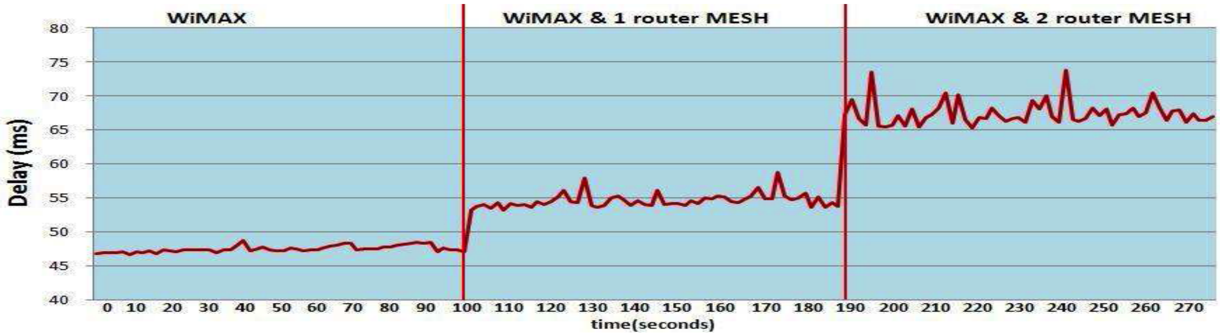


Figure 3. Average Delay (msecs)

Fig. 2 shows the maximum average throughput of each test (in Kbits per second), that can be reached without impairment of the streaming video. For the first test, we obtained a throughput of 37 Mbps with a few fluctuations in WiMAX. For the second one, we have got 17 Mbps and 6 Mbps in third one. All of them allow a quality of video higher than the demanded.

The average delay of the messages is shown at Fig. 3. It shows the performance of using different channels on 802.11 networks, as well as the achievement of requirements, since this value should not be higher to 150 ms, being 75 ms approximately in a worst scenario. In addition to the results showed in the aforementioned graphs, we have also obtained a 9% of average packet lost ratio at worst case, which is greater than the requirements of the application.

VII. CONCLUSIONS AND FUTURE WORK

The application of telecommunication technologies in areas like first response and emergency services is essential to be able to provide faster responses under critical situations or natural disasters. The networking infrastructure presented in the paper efficiently covers the requirements of applications located in this frame, and the carried out field trials study the performance of the system.

According to collected results, the efficiency of both the mesh routers and the WiMAX deployment is suitable for this type of multimedia services. This scenario comprises a simple test bed, but it effectively offer support to services demanded by the staff of emergency services.

The analysis in this paper is found very useful to identify the major problems and benefits of both WiMAX and WMNs in the development of an emergency management platform. The paper presents a real solution tested in cooperation with emergency services of Region de Murcia. A great effort is currently being carried out to unify those platforms and projects orientated to emergency services, and solve several technical problems involving wireless technologies. The aim of this research line is develop an integral infrastructure capable of hosting many emergency services, especially oriented to manage disaster conditions.

ACKNOWLEDGEMENTS

Part of this work has been funded by the excellence research groups funds of Fundacion Seneca under grant 04552/GERM/06.

This work has been also partially funded by excellence research group funds from the Spanish CARM and project TIN2008-06441-C02-02.

REFERENCES

[1] M.Protmann, Asad A. Pirzada, "Wireless Mesh Networks for Public Safety and Crisis Management Applications", National ICT Australia, Feb 2008

[2] Raheleh B. Dilmaghani and Ramesh R. Rao. "Hybrid Wireless Mesh Network with Application to Emergency Scenarios", Journal of Software, vol 3, no. 2, Feb 2008

[3] S. Bouckaert, J. Bergs, D. Naudts, "A Mobile Crisis Management System for Emergency Services: from Concept to Field Test", Qshine 2006, Ontario.

[4] K. Balachandran et al., "Mobile Responder Communication Networks for Public Safety" IEEE Comm. Magazine, vol. 44 no.1, 2006, pp.56-64.

[5] Yan Zhang, Jijun Luo ,Honglin Hu, "Wireless Mesh Networking", 2007 Cap 15, pp.512-513.

[6] C. Gomez, M. Catalan, X.Mantecon and J. Paradells. "Evaluating Performance of Real Ad-hoc Networks Using AODV with Hello Message Mechanism for Maintaining Local Connectivity", PIMRC 2005, Berlin.

[7] SAFECOM Programs, "Statement of Requirements for Public Safety Wireless Communication and Interoperability", US Dept. of Homeland Security.

[8] R. Draves, J. Padhye and B. Zill. "Routing in Multi-radio, Multi-hop Wireless Mesh Networks", ACM MobiCom, Philadelphia, PA, September 2004.

[9] Francesco Chiti, R. Fantacci, L. Maccari, D. Marabissi and D. Tarchi, "A Broadband Wireless Communications System for Emergency Management", IEEE Wireless Communications, June 2008.

[10] R. Fantacci and D. Tarchi, "Bridging Solutions for a Heterogeneous WiMAX-WiFi Scenario", J. Comm. and Networks, vol 8, no.4, 2006, pp.369-377

[11] T. Clausen and P. Jacquet. Optimized link state routing protocol (OLSR). RFC 3626, www.olsr.org

F-REX: Event Driven Synchronized Multimedia Model Visualization

Dennis Andersson
Swedish Defense Research Agency
dennis.andersson@foi.se

Abstract

Reconstruction and Exploration (R&E) was developed to analyze complex chains of events in distributed tactical operations. The approach specifically points out domain analysis, modeling, instrumentation and data collection as the reconstruction steps that will enable exploration through presentation. In reality however, analysts often want to iterate the presentation step and feed back data into the model enabling iterative analysis. This work presents an improved version of the R&E approach that better fits the way analysts work.

While it would be possible to force the improved version of R&E into existing tools, the increasing amount of multimedia data becoming available, such as video and audio, motivates a redesign of existing tools to better support the new model. This paper also presents F-REX as the first tool tailored to deal with multimedia rich models for R&E and streamlined to follow the improved R&E approach.

1. Introduction

Analyzing cause and effect in a complex chain of events spanning over a large area is a very difficult task for any analyst since the analyst will need to understand what is going on at multiple locations simultaneously. It is obviously impossible for an analyst to observe everything first hand, methods and tools are needed to overcome this problem. One promising approach is Reconstruction & Exploration (R&E) [8] that makes use of a multimedia model of the operation and enables post action analysis. R&E has been used successfully in several domains, such as military exercises, live fire brigade operations, staff exercises and more.

Closely linked to R&E is the MIND framework, which was the reference implementation of a toolset supporting R&E [7], [8]. This system is streamlined to support modeling, instrumentation, data collection and presentation and makes for an excellent tool to support debriefings or after-action reviews (AARs) [9], [10], [6]. However, after several years' usage it has become more and more apparent that the design of MIND does not scale very well to the increasing amount of data that becomes available as technology becomes more sophisticated. Also the R&E model does not capture the way analysts work in practice, so the updating of MIND also called for an update of the R&E model to fill in the gaps to better reflect how it is being used in practice. The improved R&E approach is in this paper referred to as the F-REX approach [3] to distinguish between the two versions. The new tool, F-REX Studio, is streamlined to fit the F-REX approach.

2. Design Goals

R&E was designed to let the analyst play back the course of events much like one would do in a DVD player for example, and then pause or stop to interact with a certain set of data when something interesting shows up in one of the data streams being presented. This method has proven easy to use for analysts even with little computer experience, albeit the procedure of assembling data and couple it to the model is more difficult and requires understanding of the underlying models. Although this is more a property of the MIND framework than it is of the actual R&E approach, it is a weakness of the approach that the approach does not capture and support this in a satisfying manner.

The main design goals for F-REX are thus to maintain ease of use for analysts and simplify the process of getting data ready for analysis and presentation. Further the approach is intended to be very general and usable in many different scenarios, ranging from strategic level down to operational level. Bearing that in mind and the fact that new technology constantly offers new alternatives for data capture in ways that are impossible to foresee, the approach should not rely on any particular data source but be

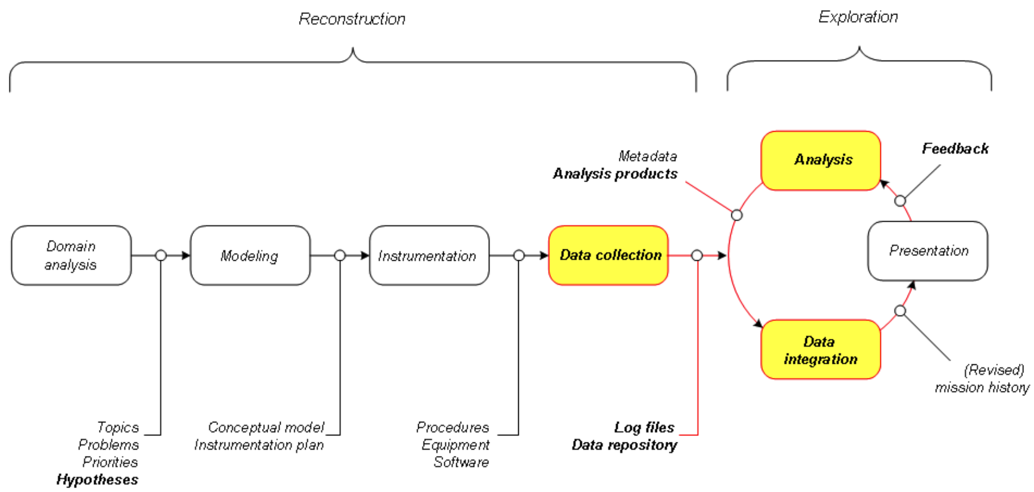


Figure 1. The improved R&E approach workflow with changes from original R&E outlined.

flexible enough to support just about any type of data coming from any source.

As for F-REX studio, this must support the F-REX approach fully and offer a platform onto which it is easy to develop new modules that make use of new data or visualization techniques as they become available. One must also bear in mind that the amount of data available for capture is very likely to continue to grow and therefore F-REX Studio should not introduce any restrictions on data capacity. The final design goal that was defined is the ability to easily cooperate between analysts so that multiple analysts can simultaneously work on the same dataset. Again this is not something directly restricted in R&E, but the lack of its explicit support explains why it has not been implemented in MIND.

To sum up, the most important design goals for F-REX and F-REX Studio are flexibility, scalability, cooperability, extensibility and usability.

3. The F-REX Approach

The R&E approach [8] is commonly described as a process leading from domain analysis to presentation via modeling, instrumentation and data collection. This same description serves as a base for the F-REX improvements of R&E. The new features in the F-REX approach are highlighted in Figure 1.

The domain analysis, modeling and instrumentation phases remain virtually unchanged from their original definition in R&E. The data collection step, however, has been split into Data collection and Data integration. The Data collection phase is the phase where the actual data is automatically captured or manually collected, according to the plans defined in the Instrumentation phase. Data may consist of scribbled notes, system log files, photographs, multimedia feeds or any other available data. The data integration phase serves to integrate the captured data with the conceptual model and couple it to the research questions defined in the domain analysis. This data coupling prepares the model for playback by categorizing, sorting and coding data as necessary.

The presentation phase is the final phase of R&E. During the presentation phase the model is played back from the start to the end and a set of data visualizers are updated as the chain of events unfolds. This allows the audience, i.e. at an AAR, to relive the operation and see what happens at different locations during the entire operation, giving the analyst a chance to relate individual actions to the global picture and draw conclusions that would be impossible from traditional observation on a single location. This enables the analyst to detect anomalies from the expected course of events and other data of particular interest. R&E does not separate presentation from analysis and it is unclear what the end product of the Presentation really is. The F-REX approach tries to remedy this by stating that the presentation step serves a mean to detect interesting data that the analyst may want to investigate further.

The analysts will use the presentation feedback as a starting point for their analysis and then dig deeper into the model to try to answer questions or hypotheses. In the case of abstract questions or complex relations between events, parts of the analysis results may be integrated into the model again to enrich the model for a new presentation and new analysis. This turns the Exploration phase into a loop, which will continue until the problems and hypotheses have been properly investigated.

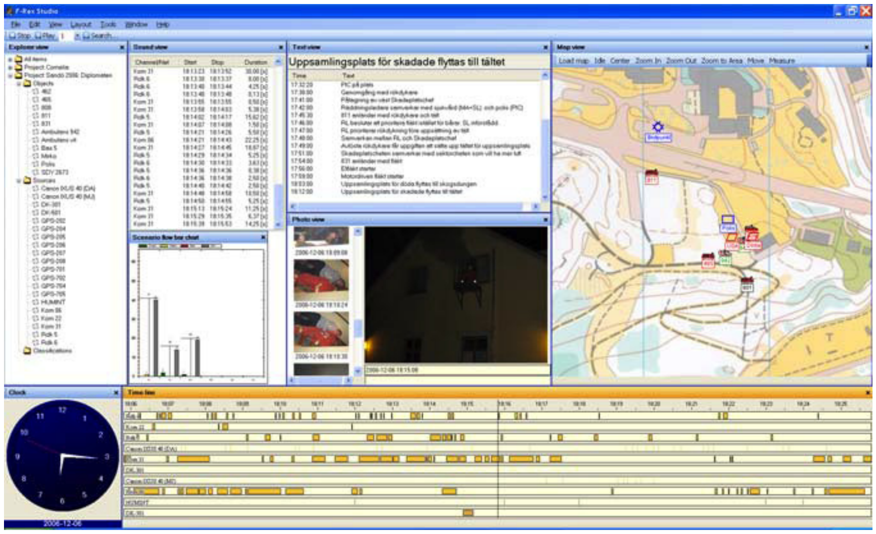


Figure 2. Screenshot of F-REX Studio showing one layout, presenting multimedia, observer notes, statistics and GIS information from a rescue services exercise in northern Sweden 2006.

4. F-REX Software

The main software that has been designed is the F-REX Studio which replaces MIND for R&E as the main engine for modeling and presentation (Figure 2). A wide range of standalone recording and conversion tools have been developed to support data capture, as well as applications to control and monitor data capture remotely via a network where available.

Data integration is fully integrated into the Studio and extensions for instrumentation are being planned alongside integration of data capture tools. It has been recognized however that it is neither possible nor desirable to fully integrate everything into the Studio, for instance standalone data capture systems like handheld cameras, voice recorders and proprietary systems which may be more practical to operate standalone and instead import their data output manually afterwards. Data capture systems that can be connected to a F-REX server in some way, such as NBOT [14] or any network enabled software, may however benefit from being directly integrated into the Studio to allow for automation of the otherwise labor intense data integration process.

All data that is imported is automatically synchronized using timestamps from the recorders. However, these timestamps have proven not very trustworthy due to drifting clocks. Therefore F-REX supports a multitude of ways to resynchronize data semi-automatically or manually depending on the complexity of the clock drifts.

Analysis is partly supported by the F-REX Studio. Some custom analysis tools for certain types of detailed analysis have been built in, and the framework allows for addition of more tools as they become necessary.

5. F-REX Studio Architecture

The F-REX Studio is built as a desktop application with loosely coupled modules that can communicate with each other and the framework in a standardized manner. The main framework architecture is typically envisaged as a multi-tier architecture [4] with a clear distinction between the four defined tiers (Figure 3). The framework implements the tiers and provides access to basic routines and common visualization features. Each module on the other hand is implemented according to the Model View Controller paradigm (MVC) [11]. The model in this case is provided by the framework while the view and the controller are programmed by the module developer, assisted by the common routines and definitions available in the framework.

One of the main reasons for using a 4-tier architecture is the ability to separate the data repository from the implementation to allow for a modification of the physical data structure without having to change the main code. The 4 tiers in the model are therefore defined as data tier, data access tier, business tier and presentation tier.

5.1. Data Tier

The data tier provides the data storage. What storage facility to use can be configured at runtime by the user in F-REX Studio. Several experimental solutions have been tested briefly with object oriented

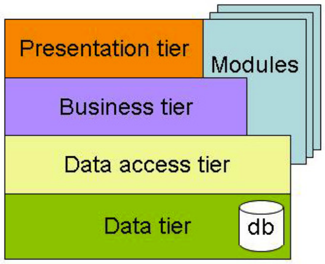


Figure 3. F-REX Studio modeled as a 4-tier architecture with plugin modules interfacing the top three tiers.

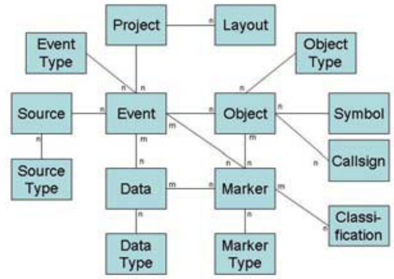


Figure 4. The main entities and their relations in the F-REX data tier.

databases and file based solutions, however the preferred solution, that is also mostly used, is based on a relational database (Figure 4).

The most central entities in the data tier are *Events*, *Data* and *Objects*. An *Event* represents the occurrence of new *Data*, for example a new photo available from a certain camera (represented as a *Source*). The *Event* entity contains time, duration and type. The entity will typically be linked to one or more *Data* entities containing any type of *Data* related to the *Event*, for instance photo, position, comment or camera settings. Further, the *Event* can be linked to any number of *Objects*, representing for instance the photographer or the photo subject. Coupling data in this way allows for automatic processing and filtering of data to quickly extract useful information.

5.2. Data Access Tier

The data access tier defines the interfaces that are implemented by the data tier. These interfaces are accessed by the modules and business tier, allowing uniform access to the data regardless of the implementation used in the data tier. Since all access to the data tier is routed via this tier, different filtering and other useful data manipulation procedures can effectively be handled by the data access tier.

5.3. Business Tier

The business tier is split into two parts, the base services and the plugin support modules. The base services provide the main event engine that for instance makes sure all modules are synchronized and loaded with the right data at the right time. It also provides a useful message passing scheme to allow the modules to communicate with each other.

The plugin support modules are basically helper classes and interfaces that assist the programmer in developing plugins that will inherit the F-REX look and feel. Although they are not required they will help the programmer to quickly get access to the data and functionality supplied by the base services.

5.4. Presentation Tier

All user interfaces are located in the presentation tier. The framework provides a main workspace and docking system in which the plugin module user interfaces will reside. The framework also provides common resources and a menu system with hooks, onto which plugins can attach their own menu items.

5.5. MVC Module Architecture

A typical plugin for F-REX provides a visualizer for a certain type of data and/or events. Existing plugins have typically been developed according to the MVC architecture, with a triangular communication pattern between the model, view and controller. The modules are thus interfacing all the top three tiers of the main architecture (Figure 3).

By using the supplied base modules, the developer is given sort of a sandbox in which to develop a module, where all this is needed is to define a controller that specifies what type of events are to be supplied from the model to the view. The view can then be defined as a user control and the user interface set up as the developer prefers, and everything else will be tendered for automatically by the support modules. A solution like this has proven very useful for rapid development of new plugin modules.

The most basic plugins that have been implemented include clock displays, timeline, image, audio, video and GIS. All of these plugins contains views that are updated by the controller to synchronize against the engine clock. Among the more specialized plugins are the bookmark plugin that allows an analyst to save the current state of all visualizers and write a comment that will be tied to the current state. These “bookmarks” are automatically stored and can easily be returned to at a later stage.

The communication plugin is also worth mentioning as it gives a visual presentation of communication in a network of senders and receivers. The communication plugin was originally designed for radio communication, but the flexibility of the data tier has allowed it to successfully be used also for instance for e-mail conversations and IP communication.

6. Usage

The F-REX approach and tools have been used to successfully evaluate several exercises, for instance tactical army drills [5], strategic HQ staff exercises [12] and rescue services commander training [13]. As a proof of concept, an evaluation of professional football has also been investigated [1].

The F-REX tools and studio have mostly been used for AAR support and post mission analysis (PMA). When supporting AARs, the system has been operated by system experts that are familiar with the tools and methods. The operator assists the AAR facilitator who uses the F-REX presentation to show the participants what has happened and use this to support the discussions. This has often helped to raise the discussions from “what happened” to “why did it happen” which is a significant step forward and has been appreciated by AAR facilitators.

For PMA, analysts have typically worked in small groups, analyzing data in more or less traditional qualitative or quantitative methods using F-REX as a way to navigate through the massive datasets.

In operative work, the F-REX Studio has been used by the Swedish Police to synchronize outputs from surveillance cameras in order to match images and identify suspects. The predecessor, MIND, has also been used by the Swedish Rescue Services Agency to document live operations for feedback and analysis.

Another way of using the F-REX tools is to provide pre-action presentations (PAP) [16], which is similar to an AAR, but the audience is shown a previous exercise or operation and may reflect on events which may give them an advantage when similar situations occur in their upcoming operation.

7. Future Work

A strength, and at the same time weakness, of the F-REX approach is the tremendous amount of data that is typically collected during the data capture phase. This leads to a substantial work to integrate the data with the model before presentation can begin. For the method to be useful in an AAR context, the presentation should be done shortly after the exercise is finished. Due to the massive amount of labor needed to manually sort and integrate data, the presentation is not always as complete as would be preferred. If the infrastructure allows it, data integration should therefore be automated as much as possible so that captured data is directly integrated into the model.

Automatic data integration enables another interesting adoption of the F-REX Studio, namely live presentation of data. This would in effect make F-REX Studio a decision support system that could be integrated into a command & control (C^2) system.

The roadmap ahead also includes instrumentation support for F-REX Studio that would automatically prepare the data integration system to couple incoming data in accordance to the instrumentation plan. Partial integration of existing tools for control and monitoring of data is also planned. With these two additions the F-REX Studio would support all steps of the F-REX approach and thus become a complete F-REX system.

More plugin modules are also planned, tailored for visualizing and analyzing, for example communication in a structured manner using the extended Attribute Explorer technique [2], [15] or simple tools to organize and classify events. Other new modules being discussed are 3-D visualization, health monitoring, signal analysis, and also visualization of data flow and system communication.

Future work on the F-REX approach includes identification of compatible analysis methods to use and specifying how F-REX fits into the overall scheme of these methods. Measuring the cost and amount of time needed for high quality analysis and compare this to traditional methods is another important task that is needed to estimate how useful the F-REX approach is.

8. Conclusions

This paper presents a slight improvement of the model for the R&E approach that better maps onto how researchers and analysts work with massive multimedia intensive datasets. This model helped in defining a new framework and tool, the F-REX Studio, also described in this paper. The F-REX method and tools have been successfully used to assist multimedia intense presentations and analyses such as after action reviews and post mission analysis in several exercises and some live operations.

Flexibility is reached through the general definition of instrumentation and data capture that allows for any instruments to be used and any data to be captured. Of course this puts high demands on the F-REX Studio to be flexible in visualization. This is reached through the plugin interface which allows developers to quickly create new plugins for F-REX visualizing any data in any way imaginable automatically synchronized with all other views.

Scalability is achieved through the 4-tier architecture which allows the data access and data tiers to be exchanged for larger data warehouses, using optimized techniques to access relevant data should it be necessary. For now however, a relational database is used as the backbone which provides enough performance for the time being.

Cooperability can be reached by using a central resource for the data tier, for instance a network enabled database, allowing several analysts to work on the same set of data simultaneously.

Extensibility comes from the modular design in the business tier which allows the programmer to quickly develop new visualization modules and link to the rest of the framework to add new visualization and analysis possibilities.

Usability is mainly a feature of the presentation tier. It is up to the developer to create the user interface for any plugin modules. Using common resources helps the developer to get a common look and feel of the modules. The overall usability of the system has not yet been measured and no conclusions can be made about it so far.

9. References

[1] Albinsson, P-A. and Andersson, D., “Computer-aided football training exploiting advances in distributed tactical operations research”, *Sixth International Conference of the International Sports Engineering Association, (*Munich, Germany), Springer, New York, 2006, pp. 185-190.

[2] Albinsson, P-A. and Andersson, D., “Extending the attribute explorer to support professional team-sport analysis,” *Information Visualization 7*, Palgrave journals, doi:10.1057/palgrave.ivs.9500178, 2008, pp. 163-169.

[3] Andersson, D., Pilemalm, S. and Hallberg, N., “Evaluation of crisis management operations using Reconstruction and Exploration”, *Proceedings of the 5th International ISCRAM Conference*, Washington, DC, USA, 2008.

[4] Eckerson, W.W. "Three Tier Client/Server Architecture: Achieving Scalability, Performance, and Efficiency in Client Server Applications", *Open Information Systems,* January 1995.

[5] Hasewinkel, H. and Thorstensson, M., *OMF of air mobile battalion during Combined Challange-2006* (in Swedish), Base data report, FOI-R--1982--SE. Swedish Defence Research Agency, Linköping, Sweden, 2006.

[6] Headquarters Department of the Army, *A Leader's Guide to After-Action Reviews (TC 25-20)*, Washington, DC, 30 September 1993.

[7] Jenvald, J., *Methods and Tools in Computer-Supported Taskforce Training*, Linköping Studies in Science and Technology, Dissertation No. 598, Linköping University, Linköping, Sweden, 1999.

[8] Morin, M., *Multimedia Representation of Distributed Tactical Operations*, Linköping Studies in Science and Technology, Dissertation No. 771, Linköping University, Linköping, Sweden, 2002.

[9] Morrison, J.E. and Meliza, L. L., *Foundations of the After Action Review Process*, IDA Document 2332, Institute for Defense Analyses, Alexandria, VA, USA, DTIC/NTIS AD-A368 651, 1999.

[10] Rankin, W. J., Gentner, F.C. and Crissey, M. J., “After action review and debriefing methods: technique and technology”, *Proceedings of the 17th Interservice/Industry Training Systems and Education Conference*, Albuquerque, NM, USA, 1995.

[11] Reenskaug T., “Models – Views – Controllers” Technical Note, Xerox Parc, 1979.

[12] Thorstensson, M., Albinsson, P.-A., Johansson, M. and Andersson, D., *MARULK 2006—Methods for developing functions, units and systems*, User Report FOI-R--2188--SE, Swedish Defence Research Agency, Linköping, Sweden, 2006.

[13] Thorstensson, M., Johansson, M., Andersson, D. and Albinsson, P-A., *Improved outcome of exercises—Methods and tools for training and evaluation at the Swedish Rescue Services school at Sandö*. User Report FOI-R--2305--SE, Swedish Defence Research Agency, Linköping, Sweden, 2007.

[14] Thorstensson, M., *Using Observers for Model Based Data Collection in Distributed Tactical Operations*, Linköping Studies in Science and Technology, Thesis No. 1386, Linköping, Sweden: Linköpings universitet, 2008.

[15] Spence, R. and Tweedie, L., “The attribute explorer: information synthesis via exploration”, *Interacting with Computers 11*, 1998, pp. 137-146.

[16] Wikberg, P., Albinsson, P-A., Andersson, D., Danielsson, T., Holmström, H., Johansson, M., Thorstensson, M. and Wulff, M-E., *Methodological tools and procedures for experimentation in C2 system development - Concept development and experimentation in theory and practice*, Scientific report, FOI-R--1773--SE, Swedish Defence Research Agency, Linköping, Sweden, 2005.

Towards Integration of Different Media in a Service-Oriented Architecture for Crisis Management

Magnus Ingmarsson
Dept. of Comp. and Inform. Sci.
Linköping University
SE-581 83 Linköping, Sweden
Email: magin@ida.liu.se

Henrik Eriksson
Dept. of Comp. and Inform. Sci.
Linköping University
SE-581 83 Linköping, Sweden
Email: her@ida.liu.se

Niklas Hallberg
FOI Swedish Defence Research Agency
Olaus Magnus v. 42
SE-581 11 Linköping, Sweden
Email: niklas.hallberg@foi.se

***Abstract*—Crisis management is a complex task that involves interorganizational cooperation, sharing of information, as well as allocation and coordination of available resources and services. It is especially challenging to incorporate new, perhaps temporary, actors into the crisis-management organization while continuing to use the same command-and-control (C2) system. Based on a preceding requirement-analysis study involving interviews and workshops with crisis-management staff, we have developed a prototype C2 system that facilitates communication, collaboration, and coordination at the local-community level. A salient feature of this system is that it takes advantage of a mash-up of existing technologies, such as web-based mapping services, integrated in a open service-oriented architecture. By taking advantage of light-weight solutions capable of running as web applications within standard web browsers, it was possible to develop a scalable structure that supports decision making at multiple levels (operational to tactical) without the need to modify the system for each level. The use of C2 systems implemented as web applications creates new possibilities for incorporation of multimedia components, such as popular web-based multimedia features. In addition, we discuss the possibility of automatically integrating multimedia services into the C2 system via a service-discovery mechanism, which uses knowledge about the services and the situation to determine which services to display.**

I. Introduction

Crisis management at the local-community level is challenging in many ways [1]. Two of the most significant challenges are: (1) The management and coordination of external actors with regards to participation in solving the crisis situation and (2) the design and use of the command and control (C2) system for handling daily activities as well as extreme events. The first challenge is commonly handled by using human actors as intermediaries between the crisis-management system and the crisis-management staff. Typically, the second challenge is addressed by employing dedicated C2 systems for crisis situations.

A disadvantage of employing dedicated C2 systems, however, is that they are used in serious situations exclusively, which means relatively infrequent use. Infrequent use leads to uncertainty among the operators of how to perform certain actions within the system, which affects overall crisis-response performance. Furthermore, infrequent use contributes to a lack of knowledge about how the systems perform in real situations.

In crisis situations, time is a critical factor. Frequently, it is the case that different C2 systems as well as other information systems must interact on an ad-hoc basis. Often, these systems cannot interchange data or interpret data that other systems provide. In practice, these inabilities are currently handled by humans intermediaries and liaison staff between the crisis-management organization and the systems employed by the external actors. For example, if the crisis-management organization needs transportation, the staff is forced to contact the transportation companies directly by telephone, since the crisis-management organization does not have direct access to, or knowledge about, the systems employed by the transportation companies and the transportation resources currently available [2]. This type of ad hoc communication is sometimes a bottleneck because it draws personnel resources.

Although C2 systems can assist response commanders in situation awareness, planning, and resource allocation [3], the traditional approach to C2 systems may lead to extensive system-development times as well as difficulties in integrating the different actors and their heterogeneous systems. Unless system designers have a substantial comprehension of the different actors involved as well as their objectives, activities and information needs, the result will be systems ill-suited to the task. Furthermore, it is essential that the different actors in the local community can synchronize, coordinate, and distribute resources [4]. Moreover, it is important to integrate local and regional resources from, for example, fire and rescue services, police force, and medical-care services in the overall crisis response. Today, it is possible to develop lightweight C2 systems that facilitates cooperation based on state-of-the-art web technologies. Such web applications can integrate new services, including multimedia, in novel ways. For instance, C2 systems implemented as web applications can relatively easily support extensions consisting of a mash-up of web components from different sources.

II. Multimedia and crisis management

Although the aforementioned challenges (such as the cooperation between different actors) are significant, the incorporation of multimedia into C2 systems may help in addressing them. However, the incorporation of multimedia in traditional C2 systems have been challenging and difficult. This obstacle is particularly problematic because situational awareness is essential to crisis management.

To create C2 systems that work in real situations, it is necessary to incorporate grounded theory. A common theory used in planning for this type of situation is the OODA loop[1] [5], see Figure 1. The OODA loop states that there are different phases in the decision-making process. These phases are: Observe, Orient, Decide, and Act. To achieve a successful outcome from the decision-making process, it is important to support the different phases properly. To provide this support, the C2 system used by the commanding staff must be OODA-loop aware in that it supports to different phases in an integrated way. In Section VII, we discuss this need in detail and present how our model and current implementation tackle this issue. There have been many enhancements to the original OODA loop. Brehmer [6] proposed the Dynamic OODA (DOODA) loop model, which introduces what he refers to as "additional sources of delay" in the process. Examples of such types of delay are *information delay,* which is the time between actual outcome and the decision-maker being aware of it; *dead time,* which is the time between the initiation of an act and its actual start; and *time constant,* the time required to produce results.

III. Proposed solution

As described in the Section I, the challenges are system-related as well as organizational. The proposed solution is based on a combination of several existing technologies, and consists of two parts: Responsoria and Magubi. Responsoria is a prototype C2 system implemented as a web application. It is responsible for the interaction and connectivity between different services, devices, and users, once they are selected for inclusion in the situation [7]. Magubi is responsible for service/device/actor discovery and recommendation of different services/devices/actors. Sections IV and VI describe Responsoria and Magubi, respectively.

IV. The Responsoria model and implementation

To better understand the potential of our model, we present the basic concepts and ideas behind it. In its most simplified form, the Responsoria model is a Service Oriented Architecture (SOA), which uses web services as basis for the entire system.

A web-based user interface retains a desktop look-and-feel through the use of JavaScript while keeping the solution accessible through standard web browsers, see Figure 2. A proxy in the web server enables communication with other components, such as application servers in Responsoria. The pluggable structure extends to the application servers as well. Furthermore, Responsoria utilizes Enterprise Java Beans (EJBs) in the form of web services for mapping, note-taking, logging, etc. It is straightforward to expand the the system by incorporating other web services.

The application server currently employed is Glassfish 2, a Java-based application- and web server. Glassfish is, like the rest of Responsoria, open-source and benefits from the ability to run on a multitude of platforms and architectures, something which has been verified during development.

As mentioned, there is a desktop feel to the application itself. This is obtained by using the Google Web Toolkit (GWT), which enables developers to use Java syntax to program an entire web application. Through post-processing the application is transformed into a JavaScript application suitable for web browsers. In essence, the developer can program as accustomed to when programming a common desktop application, but still deploy it as a web-based application. The Responsoria client has been successfully tested on Apple OS X desktop, iPhone, MS Windows XP, Firefox, Internet Explorer, and Safari. We foresee that the system will work on most of the high-grade, hand-held machines currently available.

Since the system utilizes Jave EE web services, it facilitates their cross-platform distribution in much the same manner as the main program. The Java EE platform also comes with a host of features for portability, quality of service, and security.

V. A brief introduction to the prototype user interface

The basis for the development of the prototype user interface is the set of requirements identified by Pilemalm and Hallberg [8]. Figure 3 shows the main view in the user interface. To the left is the service selection panel (A). This panel lists the available resources, devices and services. We have incorporated different layouts for inclusion of the different resources. The first type of listing is an alphabetical one. Another type of listing is based on the order in which a specific task is carried out. A third type of listing may be based on recommendations from the service-discovery system. A menu bar is placed Immediately above the A panel. This enhances the perception of the application as a *desktop application.* This perception is especially prevalent if used in conjunction with a full-screen capable browser.

In the service panel itself, the currently selected activity is shown (B), see Figure 3. As can be seen in, there are tabs that enable the user to work with many different activities at the same time. Furthermore, as shown, the service panel itself also provides opportunity for incorporating different media and services. Figure 3 (B) illustrates how the system uses Google Maps together with a web-service tracking mobile phones.

Status information is displayed in the panel to the right, Figure 3 (C). Currently, three tabs display various information, such as *Request status, Activity status, and Task status.* One particularly important feature is the log, Figure 3 (D), which

[1] Although the OODA loop was originally designed for military situations, it is used in many other areas as well.

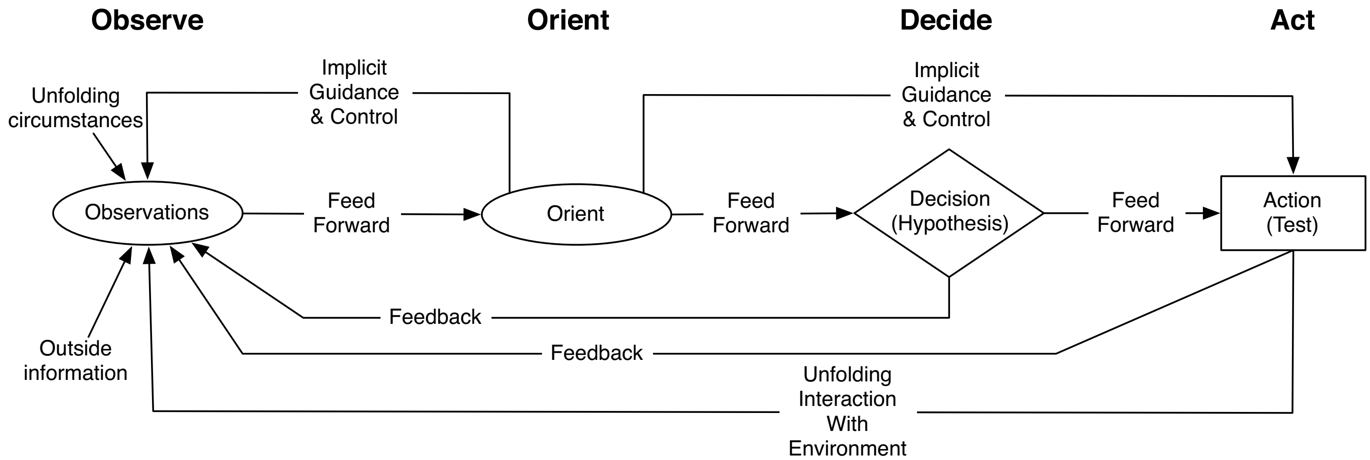


Fig. 1. The OODA loop (in Brehmer [6]). The OODA loop stands for Observe, Orient, Decide, and Act. Normally, it refers to a single person doing this cycle. However, the OODA loop can also be used when referring to organizations. Ultimately, the OODA loop describes how an individual or organization reacts to an event.

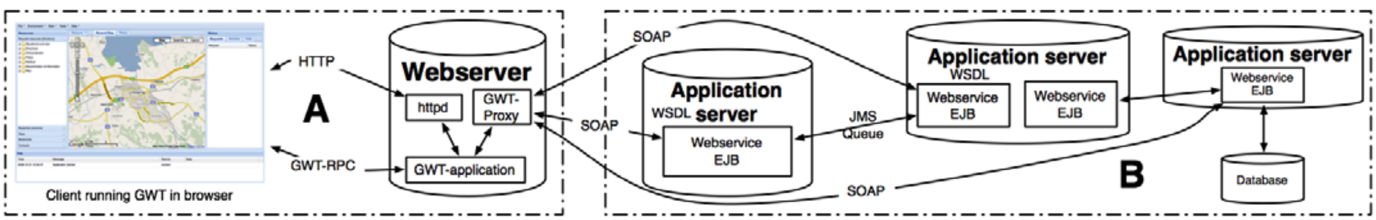


Fig. 2. Architecture of the RESPONSORIA system. (A) Web-based user-interface client. (B) Server cloud consisting of a collection of implemented web services running on application servers.

also contains a note-taking function. In order to enhance situation awareness, this log is designed to be shared by everybody using the system. It enables anyone to review what has happened, when it happened, as well as who did what.

VI. MAGUBI SERVICE DISCOVERY AND RESPONSORIA

While the RESPONSORIA model handles the usage of the services and the GUI, the MAGUBI model handles service discovery [9]. Since MAGUBI is targeted towards ubiquitous computing, it works well in crisis management situations that have many actors, services, and different types of media. Service discovery in MAGUBI can be performed in two ways:

1) *User activated.* By specifying the service or device that the user is looking for, as well as their potential properties and priorities, the user can instruct MAGUBI to search for matches.
2) *Automatic.* MAGUBI performs the service discovery itself. By using models that describe the user and world, it is able to decide which services that are of interest to the users, and subsequently execute searches proactively for them.

Figure 4 shows the MAGUBI model, which is comprised of two parts: MAGUBI and ODEN. The whole model is named MAGUBI since it is the controlling module. The two parts are surrounded by aiding modules. Starting from the bottom up in Figure 4, we can see the services and devices themselves. A peer-to-peer (P2P) subsystem keeps track of these services and devices.

ODEN is the subsystem responsible for the user-activated or more traditional service discovery. By using ontologies, ODEN is able to expand on the traditional concepts for semantic models of devices and services. After using a P2P subsystem to download semantic descriptions provided by the services and devices themselves, it evaluates them locally on the client. After evaluation the results may be presented to the user, or post-processed in the MAGUBI module.

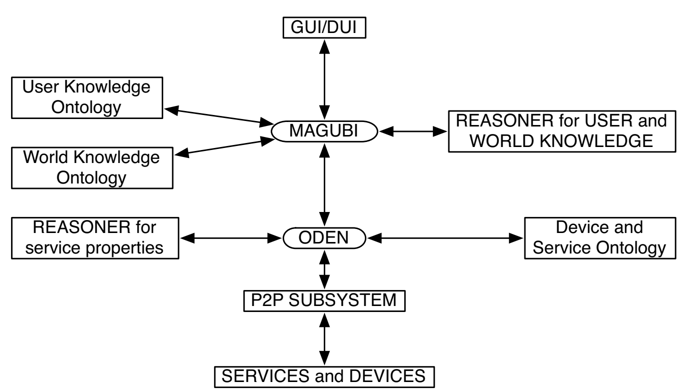


Fig. 4. The MAGUBI model.

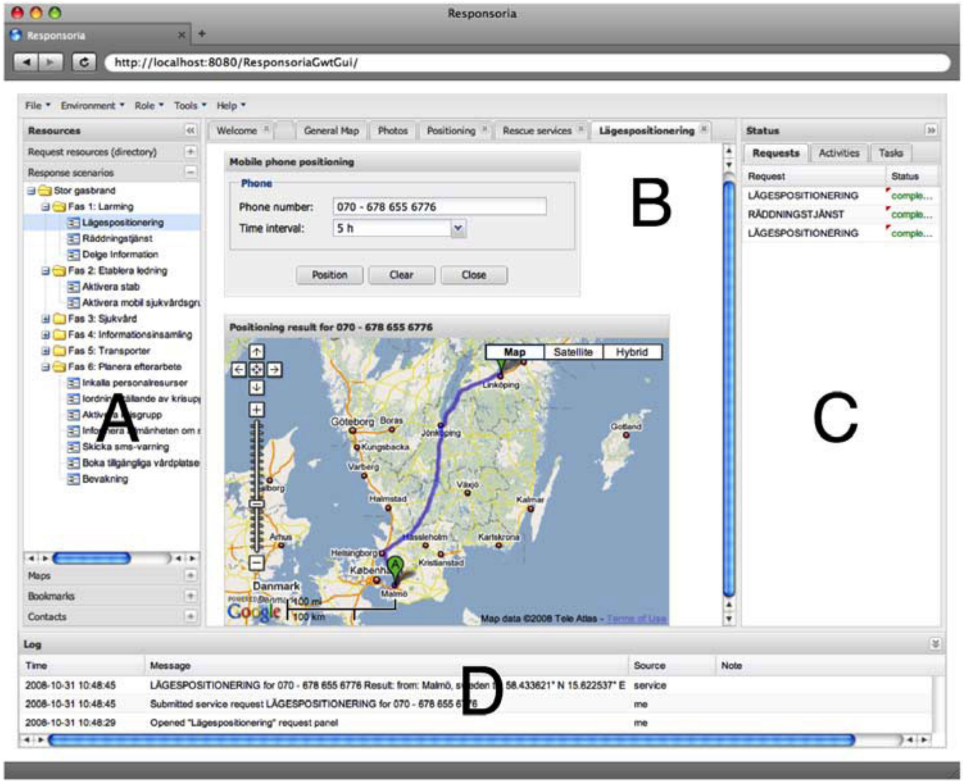


Fig. 3. RESPONSORIA main user interface: (A) Service selection panel, with preconfigured service groups for different scenarios. (B) Service panel, showing the *Mobile phone positioning service.* In this case the service panel is showing the a trace of a mobile phone for the last five hours. (C) Status panel, showing progress for different requests as well as tasks and actitivites. (D) Log, Showing all activity in the system as well as provider of a note-taking function.

The MAGUBI module may either post-process results from ODEN, or initiate searches on the users behalf. In the case of post-processing, MAGUBI inspects the results from ODEN and compares them to semantic information stored in its ontologies pertaining to: the world, devices, services, and the users. As an example the user may try and locate transportation in the form of a taxi. In an ordinary service-discovery system, the user will get a long list of available taxis. Using the ODEN subsystem, the user gets a shorter list, tailored to the exact specifications of the required properties that the user provided. MAGUBI goes one step further and may for example filter out such taxis that might very well fulfill the required transportation properties, but may soon require refueling, and as such are realistically unusable, since there is more to the transportation service than merely being able to start it.

For the proactive part, MAGUBI may commence searches for services and devices that it judges appropriate for the user. These searches are based on the user and world models, in cooperation with the rule engine and its rules. Naturally, these searches are carried out through the ODEN subsystem and are subjected to the same post-processing that user-initiated ones are.

At the head of the MAGUBI system is the GUI/DUI module [10]. In this case, it may be integrated into the web interface and accept requests from the user as well as present results from proactive searches that the MAGUBI module may do independently from the user.

VII. MULTIMEDIA TECHNOLOGIES AND THEIR INTEGRATION INTO RESPONSORIA

A. *The OODA loop and* RESPONSORIA

RESPONSORIA supports the OODA loop in multiple ways. First, it supports the first run through of the OODA loop by providing a rich environment in which to conduct observations. It is worth noting that RESPONSORIA also allows for observations to be performed from the field directly in the tool, thus supporting the *orient* part of the OODA loop. Second, it integrates tools for making sense of the observed data, such as the possibility to visualise numbers quickly as charts (see Figure 5), further aiding in the orient and decide parts of the OODA loop. Third, it provides means to effect orders onto the situation, supporting the *act* part of the OODA loop.

B. *Integration of technologies into* RESPONSORIA *through service discovery*

As mentioned above, the integration of different technologies through different services is a key factor in creating a viable crisis management system. This integration may be performed in different ways:

1) *Manual integration.* In its simplest form we are able to integrate technologies and services by just adding URLs. This may even be performed by users, either by

individually or together. In essence a wiki-type interface is created in which the users construct the application in concert.

2) *Automatic integration.* While manual integration certainly is possible, automatic is preferable. One of the most important reasons for automatic integration is the labor savings it creates. To obtain automatic integration, we propose the use of service-discovery systems such as MAGUBI.

The RESPONSORIA model, especially the web-based user interface, can benefit from both manual and the automatic integration. Today's web browsers are capable of displaying and utilizing a wide range of media and technologies out of the box. In our solution we have focused on technologies built into the browser such as JavaScript, JPG, PNG, etc. Specialized data formats may be converted using a web service.

C. Potential multimedia technologies

Since situation awareness (or orientation, as specified in OODA) is one of the highest priorities when addressing a crisis, we will briefly mention some resources and technologies that may enhance this while still being easily integrated into the main system. We will also relate these techniques to the OODA loop. By having a web-based crisis management system it is possible to rapidly tie in new services as they become available.

1) Personal video streaming: One of the possible technologies that is easy to integrate into RESPONSORIA is *live* video streaming. Services such as Bambuser [11], Qik [12], Flixwagon [13], and Kyte allow the user to broadcast *live* video over the internet using their own mobile phone as a transmitter. This means that every cellular phone is now a potential live-coverage camera in a crisis situation. This technology is instrumental in the *observe* and *orient* parts of the OODA loop.

2) Online charting: Another possible technology is charting applications, for instance the chart API from Google as shown in Figure 5, or Complan[14] as can be seen in Figure 6. It becomes easy to integrate this type of multimedia by merely including a URL. Apart from a rapid integration from URLs it is also possible to convert textual data into diagrams on the fly. These charts may be rapidly created using web services or webpages that feature simple user-interface components, such as drop down menus.

A possible drawback when using the simpler URL method is that the amount of data passed to the graphing application may cause the web server to report an error as the URL length expands beyond the web servers' limit. Nevertheless, it should be noted that the simple URL method does provide a rapid and uncomplicated way of producing charts from data.

Furthermore, the storage requirements of these diagrams are small, due to the fact that they exist as URLs. This existence by URL also has the added benefit of saving bandwidth and computing time for the crisis management center since pictures will not be served from the crisis management's own data center but from a third party.

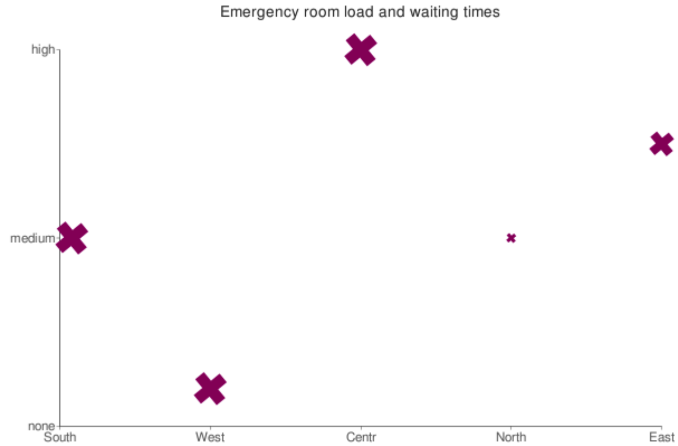


Fig. 5. Chart generated from the following URL:
http://chart.apis.google.com/chart?chxt=x,y&chtt=Emergency%20room%20load%20and%20waiting%20times&cht=s&chxl=0:|South|West|Centr|North|East|1:|none|medium|high&chs=300x200&chd=t:2,25,50,75,100|50,0,100,50,75|100,100,100,30,75&chm=x,990066,0,0,30,1
The emergency room load is shown for the different locations as none, medium, or high. The size of the X indicates relative waiting time.

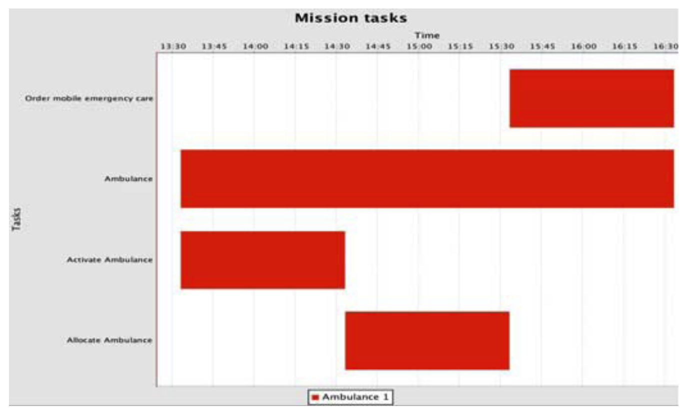


Fig. 6. Complan showing different tasks to be performed in a crisis scenario and when to do them.

With regards to the OODA loop, online charting fits in the *decision* part, since it provides supportive information regarding which direction to go.

3) Online animations: Using technologies, such as OpenLazlo, enables data from formats such as XML to be converted into for instance Flash or DHTML to be easily accessible online [15]. Ming [16] is a similar framework that generates Flash on the fly.

VIII. DISCUSSION

In this section, we discuss the redundancy feature of RESPONSORIA and the service-discovery mechanism that facilitates the integration of different technologies, media, and services.

A. Redundancy

There are several layers of redundancy in our model. As seen in part A of Figure 2, even though the web server is the weakest link in the concept, the server side can be hardened through off-the-shelf web-server technology solutions, such as backup servers that automatically engage if the main server fails and other JavaEE features [17]. Part B of Figure 2 shows an example of the application servers. It is very likely that there will be a surplus of servers offering similar if not identical services. Through the use of service-discovery such as MAGUBI, rapid recovery is ensured if services fail.

B. Service Discovery

The technical aspects of service discovery through the use of the custom-built application MAGUBI has been mentioned in Section VI. Here, we will address the non-technical part of MAGUBI, namely its philosophical underpinnings. MAGUBI is a service discovery model and implementation that addresses service discovery from the perspective of the user rather than the system. This perspective connotes an attempt to address the issue with discovery *and* selection of services.

Many traditional service-discovery systems only address the *discovery* part of service discovery. On one hand this focus helps the user since services are discovered. On the other hand it leaves the user wanting when it comes how to perform service selection since the user has to do the evaluation and selection manually. With MAGUBI, this evaluation and selection is offloaded from the user onto the service-discovery system. Furthermore, since MAGUBI has information about the world, situation, and the users, it is able to perform proactive suggestions in terms of services based on what it deems necessary at the present time.

C. Multimedia

As mentioned above, different multimedia services support different parts of the OODA loop. A service provider may aid in the configuration of an interface by providing information in the service descriptions about where in the OODA loop his/her particular service fits in. MAGUBI supports automatic classification into the different OODA categories depending on which rules are entered. By having this ability, we believe that greater efficiency is achieved regarding where in the GUI to position available services as well as supporting which services to include in the GUI in the first place.

IX. SUMMARY AND CONCLUSION

We have presented an approach to C2 systems for interorganizational cooperation at the local-community level. The prototype system is a web application, which means that it does not require client installation. The use of web browsers running on standard hardware makes it highly available to users in crisis situations. Furthermore, this approach enables relatively straightforward incorporation of rich multimedia into the C2 system as well as mash-ups of multimedia components. The incorporation of multimedia components can be done in different ways—both manually and automatically.

A service-discovery system can potentially facilitate the automatic discovery and inclusion of services by using knowledge about the situation and the services available, as well as general world information. The prospect of proactive inclusion of services and multimedia through the service-discovery system is appealing. We believe that the service-discovery view of multimedia mash-ups, combined with rapid inclusion and dismissal of actors and services, can be used to develop new types of dynamic C2 systems. Moreover, we believe that it is important for the C2 system to be aware of the general C2 method used (for instance the OODA and DOODA loops) and to provide focused support for the different stages of the decision-making process.

ACKNOWLEDGMENTS

This work has become possible due to grants from the Swedish Emergency Management Agency (KBM). We thank Ola Leifler and Jiri Trnka for valuable discussions and suggestions for improving the manuscript.

REFERENCES

[1] G. D. Haddow and J. A. Bullock, *Introduction to Emergency Management*. Butterworth-Heinemann, Boston, MA., 2006.
[2] D. Mendonça, T. Jefferson, and J. Harrald, "Collaborative adhocracies and mix-and-match technologies in emergency management," *Communications of the ACM*, vol. 50, no. 3, pp. 44–49, 2007.
[3] E. Jungert, N. Hallberg, and A. Hunstad, "A service-based command and control systems architecture for crisis management," *The International Journal of Emergency Management*, vol. 3, no. 2, pp. 131–148, 2006.
[4] S. Y. Shen and M. J. Shaw, "Managing coordination in emergency response systems with information technologies," in *Proceedings of the Tenth Americas Conference on Information Systems*, New York, NY, USA, 2004.
[5] G. T. Hammond, *The Mind of War: John Boyd and American Security*. Washington D.C., U.S.A.: Smithsonian Institution Press, 2001.
[6] B. Brehmer, "The dynamic ooda loop: Amalgamating boyd's ooda loop and the cybernetic approach to command and control," in *10 th International Command and Control Research and Technology Symposium*, McLean, Virginia, U.S.A., 2005.
[7] M. Ingmarsson, H. Eriksson, and N. Hallberg, "Exploring development of service-oriented c2 systems for emergency response," in *Proceedings of the 6th International ISCRAM Conference – Gothenburg, Sweden*, J. Landgren, U. Nulden, and B. V. de Walle, Eds., May 2009.
[8] S. Pilemalm and N. Hallberg, "Exploring service-oriented c2 support for emergency response for local communities," in *Proceedings of ISCRAM 2008*, Washington DC, 2008.
[9] M. Ingmarsson, *Modelling User Tasks and Intentions for Service Discovery in Ubiquitous Computing*. Ph. Lic. Thesis, Linköpings universitet, 2007.
[10] A. Larsson and M. Ingmarsson, "Ubiquitous information access through distributed user interfaces and ontology based service discovery," in *Multi-User and Ubiquitous User Interfaces at MU3I-06*, A. Butz, C. Kray, A. Krüger, and C. Schwesig, Eds., 2006.
[11] (2009, 03). [Online]. Available: http://bambuser.com/
[12] (2009, 03). [Online]. Available: http://qik.com/
[13] (2009, 03). [Online]. Available: http://www.flixwagon.com/
[14] O. Leifler, "Combining Technical and Human-Centered Strategies for Decision Support in Command and Control — The ComPlan Approach," in *Proceedings of the 5th International Conference on Information Systems for Crisis Response and Management*, May 2008.
[15] (2009, 04). [Online]. Available: http://www.openlaszlo.org/
[16] (2009, 04). [Online]. Available: http://ming.sourceforge.net/
[17] (2009, 03). [Online]. Available: http://java.sun.com/javaee/

An Analysis of Two Cooperative Caching Techniques for Streaming Media in Residential Neighborhoods

Shahram Ghandeharizadeh, Shahin Shayandeh, Yasser Altowim
shahram@usc.edu, shayande@usc.edu, altowim@usc.edu
Computer Science Department
University of Southern California
Los Angeles, California 90089

Abstract

Domical is a recently introduced cooperative caching technique for streaming media (audio and video clips) in wireless home networks. It employs asymmetry of the available link bandwidths to control placement of data across the caches of different devices. A key research question is what are the merits of this design decision. To answer this question, we compare Domical with DCOORD, a cooperative caching technique that ignores asymmetry of network link bandwidths in its caching decisions. We perform a qualitative and quantitative analysis of these two techniques. The quantitative analysis focuses on startup latency defined as the delay incurred from when a device references a clip to the onset of its display. Obtained results show Domical enhances this metric significantly when compared with DCOORD inside a wireless home network. The qualitative analysis shows DCOORD is a scalable technique that is appropriate for networks consisting of many devices. While Domical is not appropriate for such networks, we do not anticipate a home network to exceed more than a handful of wireless devices.

1. Introduction

Advances in mass-storage, networking, and computing have made streaming of continuous media, audio and video clips, in residential neighborhoods feasible. Today, the last-mile limitation has been resolved using a variety of wired solutions such as Cable, DSL, and fiber. Inside the home, computers and consumer electronic devices have converged to offer plug-n-play devices without wires. It is not uncommon to find a Plasma TV with wireless connectivity to a DVD player, a time shifted programming device (DVR) such as Tivo, a cable set-top box, a game console such as Xbox, and a computer or a laptop. The primary constraint of this home network[1] is the radio range of devices and the available network bandwidth connecting devices.

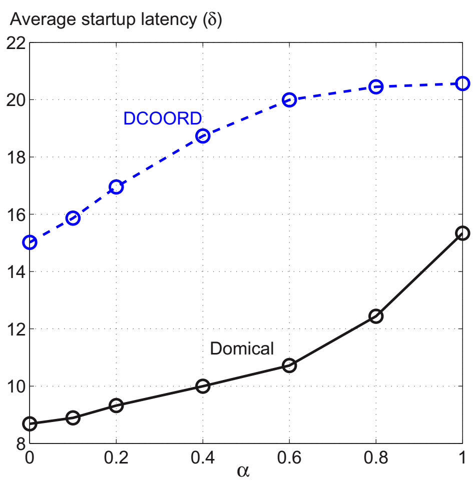
The wireless in-home networks are attributed to consumer demand for no wires, ease of deploying a wireless network, and the inexpensive plug-n-play components that convert existing wired devices into wireless ones. A device might be configured with an inexpensive[2] magnetic disk drive and provide hybrid functionalities. For example, a cable box might be accompanied with a magnetic disk drive and provide DVR functionalities [8]. A device may use its storage to cache content.

DCOORD [1] and Domical [5] are two cooperative caching techniques for residential neighborhoods. While DCOORD is designed for home gateways in a neighborhood, Domical targets devices inside the wireless home. A qualitative comparison of these two techniques is shown in Table 1. This table shows DCOORD assumes abundant network bandwidth and employs a decentralized hash table to scale to hundreds and thousands of home gateways in a residential neighborhood. Domical, on the other hand, targets an in-home network consisting of a hand-full of devices.

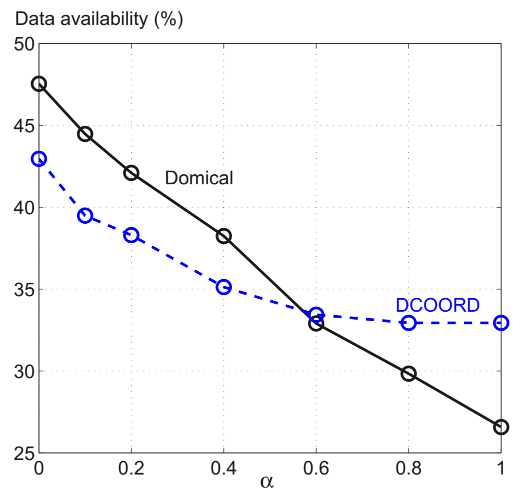
Both DCOORD and Domical partition the available storage space of a device into two areas: a) private space, and b) cache space. The private space is for use by the client's applications. Both techniques manage the cache space of participating devices and their contents. A parameter, α, controls what fraction of cache space is managed in a greedy manner. When α=0, the device is fully cooperative by contributing all of its cache space for collaboration with other devices. When α=1, the device acts greedy by using a technique such as LRU or DYNSimple [4] to enhance a local optimization metric such as cache hit rate. Both DCOORD and Domical support these extreme and intermediate α values.

[1] Power becomes a constraint when a mobile device is removed from the network for use outside the home.

[2] The cost per Gigabyte of magnetic disk is less than 10 cents for 1.5 Terabyte disk drives.



1.a) Startup latency



1.b) Data availability

Figure 1. Different α values, Domical Vs DCOORD, $UK1X$, $\mu = 0.73$, $\frac{S_T}{S_{WS}} = 0.5$

DCOORD and Domical have different objectives. While Domical strives to minimize the likelihood of bottleneck link formation in a wireless network, DCOORD strives to maximize both the cache hit rate of each node and the number of unique clips stored across the nodes of a cooperative group. In addition, their design decisions are different. DCOORD caches data at the granularity of a clip while Domical supports caching at the granularity of both clips and blocks. (Section 2 shows block caching enhances the startup latency observed with Domical.) Finally, DCOORD chooses victim objects using a recency metric while Domical considers both the frequency of access to objects and their size.

Since Domical was designed for use with a handful of devices, it may not substitute for DCOORD outside the home when the neighborhood consists of hundreds of household. This raises the following interesting question: Is it possible for DCOORD to substitute for Domical inside a wireless home? The short answer is a "No" because of the asymmetric bandwidth of the wireless links between devices. To elaborate, a recent study [7] analyzed deployment of six wireless devices in different homes in United States and England. It made two key observations. First, the bandwidth of wireless connections between devices is asymmetric. Second, this study observed that an ad hoc communication provides a higher bandwidth when compared with a deployment that employs an access point because it avoids the use of low bandwidth connection(s).

The primary contribution of this study is to quantify the merits of a cooperative caching technique such as Domical that controls placement of data across devices by considering the asymmetry of their wireless link bandwidths. We use DCOORD as a comparison yard-stick because it is the only cooperative caching technique that is comparable to Domical. Obtained results show Domical enhances startup latency observed by different devices significantly. This implies that for a wireless home network an appropriate cooperative caching technique should consider bandwidth configurations between different devices.

	DCOORD	Domical
Scalable	Yes	No
Limited Network Bandwidth	No	Yes
Employs object size	No	Yes
Data granularity	Clip	Clip/Block

Table 1. A qualitative analysis.

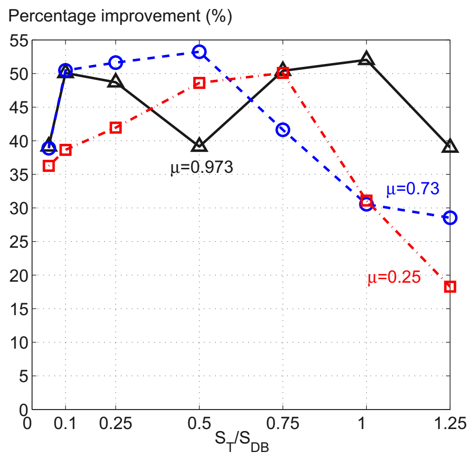
A secondary contribution is to highlight caching of data at the granularity of block when network bandwidth and storage are abundant. With a cooperative technique such as Domical, block (instead of clip) caching enhances startup latency.

To the best of our knowledge no study quantifies the performance of two different cooperative caching techniques for streaming media in a wireless home network. Due to lack of space, we have eliminated a discussion of other cooperative caching techniques and refer the interested reader to [6] for this survey.

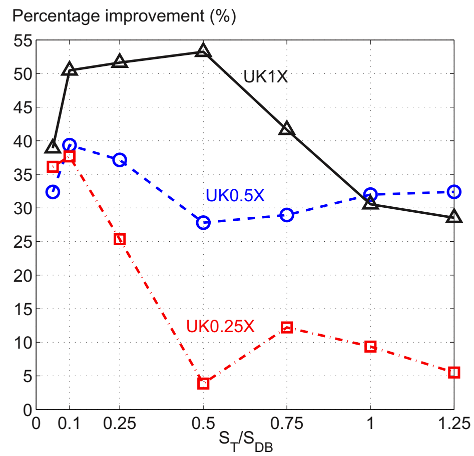
The rest of this paper is organized as follows. Section 2 provides a quantitative comparison of these two techniques. We conclude with future research directions in Section 3.

2. A simulation study

When one compares Domical and DCOORD, the following natural questions arise: Is it possible for DCOORD to



2.a) $UK1X$



2.b) $\mu = 0.73$

Figure 2. Percentage improvement in startup latency by Domical in comparison with DCOORD, α=0.

substitute for Domical? And, if Domical is better then how much better is it? To answer these questions, we built a simulation model of both DCOORD and Domical. This model assumes a household consisting of six wireless devices with wireless network bandwidths identical to those of a United Kingdom household reported in [7]. This household is denoted as UK1X. We scale down the link bandwidths by a factor of 2 and 4 to construct two hypothetical households, UK0.5X and UK0.25X.

We assumed a heterogeneous repository consisting of 864 clips. All are video clips belonging to two media types with display bandwidth requirements of 2 and 4 Mbps. The 432 clips that constitute each media type are evenly divided into those with a display time of 30, 60, and 120 minutes. The total repository size, S_{DB}, is fixed at 1.29 Terabytes. Each device is configured with the same amount of cache space and the total size of this cache in the network is S_T. In our experiments, we manipulate the value of S_T by reporting the ratio $\frac{S_T}{S_{DB}}$.

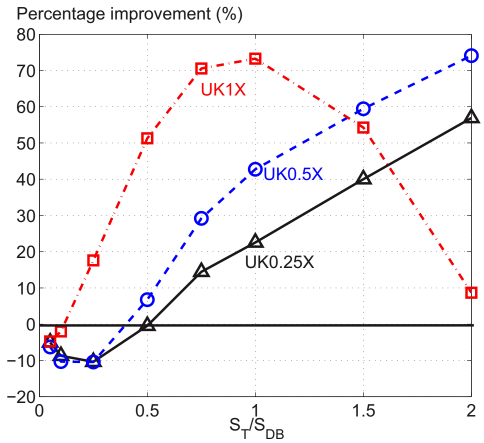
We use a Zipf-like distribution [2] with mean of μ to generate requests for different clips. One node in the system is designated to admit requests in the network by reserving link bandwidth on behalf of a stream. This node, denoted N_{admit}, implements the Ford-Fulkerson algorithm [3] to reserve link bandwidths. When there are multiple paths available, N_{admit} chooses the path to minimize startup latency.

The simulator conducts ten thousand rounds. In each round, we select nodes one at a time in a round-robbin manner, ensuring that every node has a chance to be the first to stream a clip in the network. A node (say N_1) references a clip using a random number generator conditioned by the assumed Zipf-like distribution. If this clip resides in N_1's local storage then its display incurs a zero startup latency. Otherwise, N_1 identifies those nodes containing its referenced clips, termed candidate servers. Next, it contacts N_{admit} to reserve a path from one of the candidate servers. N_{admit} provides N_1 with the amount of reserved bandwidth, the paths it must utilize, and how long it must wait prior to streaming the clip. This delay is the incurred startup latency.

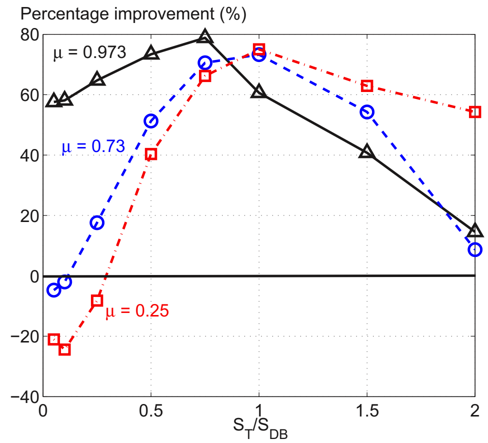
Performance results: Figure 1.a shows the average startup latency with Domical and DCOORD as a function of different α values. When compared with one another, Domical enhances average startup latency by approximately 40% to 50%. It is interesting to note that Domical results in higher availability of data for α values less than 0.6, see Figure 1.b. This means the dependencies between the caches of different devices (constructed by Domical) is effective in maximizing the number of unique clips in the home network. With $\alpha = 1$, DCOORD provides a higher availability because (a) it employs a hash function to assign clips to nodes, and (b) when a clip assigned to N_i is referenced by a neighboring device, N_i places this clip as the next to be evicted from N_i's local storage. Such a mechanism does not exist with Domical.

Domical provides a lower startup latency than DCOORD because it assigns the frequently accessed clips to the device with the highest out-going link bandwidths. This minimizes the formation of bottleneck links in the wireless network, reducing the possibility of a device waiting for an active display of a clip to end.

In almost all our experiments, Domical outperforms DCOORD. In Figure 2.a, we show the percentage improvement in startup latency observed by Domical when compared with DCOORD for different distributions of access to clips and $\alpha = 0$ using the network bandwidth observed from the UK household of [7]. In this experiment, we vary the total cache size (S_T) on the x-axis. Even with an access distribution that resembles a uniform distribution ($\mu = 0.25$), Domical outperforms DCOORD because it ma-



3.a $\mu = 0.73$



3.b $UK1X$

Figure 3. Percentage improvement with block-based caching when compared with clip-based caching using Domical.

terializes a larger number of unique clips across the cooperative cache.

The bandwidth of the wireless links has an impact on the margin of improvement provided by Domical. This is shown in Figure 2.b where we analyze the impact of scaling down wireless link bandwidths: Factor of two and four relative to the original observed link bandwidths, termed UK0.5x and UK0.25X, respectively. The percentage improvement observed by Domical drops because the bandwidth of wireless links are so low that formation of bottlenecks is very high.

One may improve the startup latencies observed with Domical by changing the granularity of caching from clip to block. This is because Domical pre-stages the first few blocks of different clips across the network strategically in order to minimize the startup latency. This is shown in Figure 3 where we report on the percentage improvement observed with block caching when compared with clip caching (for Domical). Note that when either the available cache space or bandwidth of wireless network connections is scarce (low $\frac{S_T}{S_{DB}}$ ratios in Figure 3.a with UK0.25X), caching at the granularity of a clip is the right choice. This is because, with block-based caching, the remainder of each clip referenced by every device may involve the infrastructure outside the home, exhausting the wireless network bandwidth of the home gateway.

3. Conclusion

The asymmetric and limited bandwidth of wireless connections between devices in a household make a compelling case for a cooperative caching technique such as Domical. This is because Domical assigns data to the available cache space of different devices with the objective to minimize the likelihood of bottleneck links in the network. In this paper, we did a qualitative and quantitative comparison of Domical with DCOORD. The qualitative analysis shows Domical is not a substitute for DCOORD outside the home. The quantitative analysis shows Domical enhances average startup latency significantly when compared with DCOORD inside the home.

References

[1] H. Bahn. A Shared Cache Solution for the Home Internet Gateway. *IEEE Transactions on Consumer Electronics*, 50(1):168–172, February 2004.

[2] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker. Web Caching and Zipf-like Distributions: Evidence and Implications. In *Proceedings of INFOCOM*, pages 126–134, 1999.

[3] T. Cormen, C. Leiserson, R. Rivest, and C. Stein, editors. *Introduction to Algorithms*, chapter 26.2. MIT Press, 2001.

[4] S. Ghandeharizadeh and S. Shayandeh. Greedy Cache Management Techniques for Mobile Devices. In *Proceedings of ICDE*, pages 39–48, April 2007.

[5] S. Ghandeharizadeh and S. Shayandeh. Domical Cooperative Caching: A Novel Caching Technique for Streaming Media in Wireless Home Networks. In *Proceedings of SEDE*, pages 274–279, June 2008.

[6] S. Ghandeharizadeh, S. Shayandeh, and Y. Altowim. An Analysis of Two Cooperative Caching Techniques for Streaming Media in Residential Neighborhoods. Technical Report 2009-02, USC Database Laboratory, Los Angeles, CA, 2009.

[7] K. Papagiannaki, M. Yarvis, and W. S. Conner. Experimental characterization of home wireless networks and design implications. In *Proceedings of INFOCOM*, pages 1–13, April 2006.

[8] J. R. Quain. Cable Without a Cable Box, and TV Shows Without a TV. The New York Times, Technology Section, July 26 2007.

PopCon monitoring: web application for detailed real-time database transaction monitoring

Ignas Butėnas*, Salvatore Di Guida†, Michele de Gruttola†, Vincenzo Innocente†, Antonio Pierro‡,
*Vilnius University, 3 Universiteto St, LT-01513 Vilnius, Lithuania
†CERN Geneva 23, CH-1211, Switzerland
‡INFN-Bari - Bari University, Via Orabona 4, Bari 70126, Italy

***Abstract*—The physicists who work in the CMS experiment at the CERN LHC need to access a wide range of data coming from different sources whose information is stored in different Oracle-based databases, allocated in different servers. In this scenario, the task of monitoring different databases is a crucial database administration issue, since different information may be required depending on different users' tasks such as data transfer, inspection, planning and security issues. We present here a web application based on Python web framework, AJAJ scripts and Python modules for data mining purposes.**

To customize the GUI we record traces of user interactions that are used to build use case models.

In addition the application detects errors in database transactions (for example identify any mistake made by user, application failure, unexpected network shutdown or Structured Query Language (SQL) statement error) and provides warning messages from the different users' perspectives.

I. Introduction

In the CMS experiment[1] [2], heterogeneous resources and data are put together in different Oracle-based databases, and made available to users for a variety of different applications, such as the calibration of the various subdetector components and the reconstruction of all physical quantities.

In this complex environment it is absolutely necessary to monitor Database Resources and every application which performs database transactions in order to detect faulty situations, contract violations and user-defined events.

PopCon monitoring(Populator of Condition Objects monitoring) is an Open Source web based service implemented in Python, and designed for a heterogeneous database server, that performs data transfers to provide both fabric and application monitoring.

It promotes the adoption of the Standard web technologies, service interfaces, protocols and data models.

One of the main challenges for CMS users is to monitor their own database transactions. Moreover, different types of users need different data aggregation views depending on their role. To provide a first solution for such requirements, a new group level data aggregation, based on use case models, provided by a recorded user interaction sequence, has been recently added to PopCon monitoring.

The organization of this paper is the following: section 2 presents the *PopCon tool*[3] and its main features, section 3 presents *PopCon monitoring* architecture and features, section 4 explains how *PopCon monitoring* allows users, according to their previous record user interaction, to monitor their resources and applications, finally section 5 sums up the conclusions.

II. PopCon Tool

PopCon[3] (Populator of Condition Objects tool) is an application package fully integrated in the overall CMS framework[4] intended to store, transfer and retrieve data using Oracle-Database.

A proper reconstruction of physical quantities needs data which do not come from collision events of the *CMS experiment*: these "non event" data (*Condition data*), therefore, are stored in ORACLE Databases.

The *condition data* can be roughly divided in two groups: conditions from any lower case detector system describing its state (gas values, high low voltages, magnetic field, currents and so on), and *constants of calibrations* of the single CMS sub-detector devices, mainly evaluated in the offline analysis (pedestals, offsets, noises, constants of alignment).

CMS relies on three ORACLE databases for the condition data:

- OMDS (Online Master Database System), a pure relation database hosting online condition data from the various CMS sub-detectors;
- ORCON (Offline Reconstruction Condition DB Online System), an object-oriented database hosting conditions and calibrations needed for the high level trigger and offline event reconstruction, populated using POOL-ORA[1] technology.
- ORCOFF (Offline Reconstruction Condition Database Offline System), a master copy of Orcon in the CERN network through ORACLE streaming.

Calibration and Condition data coming from the subdetectors' computers, from network devices and from different sources (databases, ASCII files, ROOT[2] files, etc.) are packed as `C++` objects and moved to the Online condition database (ORCON) via a dedicated software package called *PopCon*. The data are then automatically streamed to the offline database (ORCOFF) and become accessible in the offline network as `C++` objects. All these database transactions generate logs which are stored in tables of a dedicated account

[1]POOL is the common persistency framework for physics applications at the LHC.

[2]ROOT is an object-oriented program and library developed by CERN and designed for particle physics data analysis.

on CMS databases, so that every transaction is traceable to a single user.

Even without LHC[5] beams, expected for the autumn of this year, this mechanism was intensively and successfully used during 2008 tests with cosmic rays and now it is under further development. Up to now, 0.5 TB of data per year have been stored into the *CMS Condition Databases*.

III. PopCon monitoring architecture and features

PopCon monitoring is structured in five main components (see Figure 1):

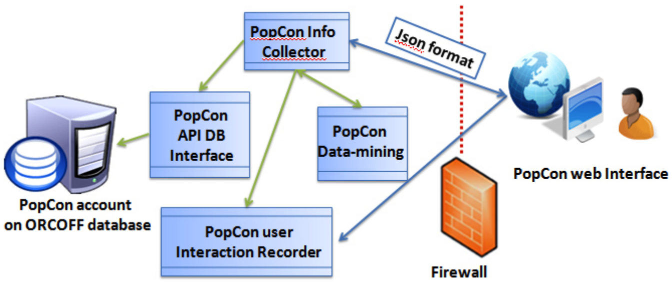


Fig. 1. PopCon monitoring Architecture

- the **PopCon API DB Interface** retrieves the entities monitored by *PopCon tool*;
- the **PopCon user Interaction Recorder** is a collection that retains an interaction history by each user.
- the **PopCon data-mining** extracts patterns from data, entities monitored by *PopCon tool* and the history of recorded user interactions, hence transforming them into information such as warnings, errors or alarms according to use case models.
- the **PopCon info collector** aggregates the information produced by the different database transactions and the history of recorded user interactions, and encodes them in JSON[3] format.
- the **PopCon Web Interface** displays the information about the database transactions from the different user perspectives, organizing data in tables (see Figure 2) and/or charts (see Figure 3).

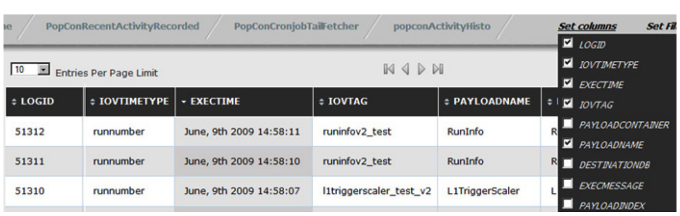


Fig. 2. The PopCon web interface represents information about database transactions in different types: both charts and tables. A user can easily add or remove columns by clicking the checkbox and also columns can be sorted. Information could be grouped according to different filters.

[3]JSON (JavaScript Object Notation) is a lightweight data-interchange format.

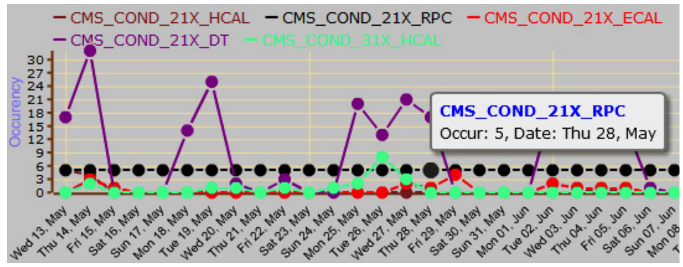


Fig. 3. PopCon Activity History: with the help of the mouse, users can interact directly with the chart (there are different types of them). Users can point the cursor to the part of chart and see the information about transactions. Charts display the accounts on which transactions were done, date and time of it and the occurences. In this picture there is an example of the linear chart.

A. PopCon API DB Interface

The *PopCon API DB Interface* is a Python script that gives access to the PopCon account on the Oracle Database. This component uses the `cx_Oracle` python module to connect to Oracle DBs and call various PL/SQL package methods.

B. PopCon user Interaction Recorder

This component creates and makes accessible the records of activities made by each user. Collected records are used to implement and improve a web interface, which can be designed for information browsing for different users in different ways. This component interacts with, and receives information from the *PopCon Web Interface*.

C. PopCon data-mining

Through the use of sophisticated algorithms this component can extract information from logs of database transactions (operator, data source, date and time, metadata) and the *PopCon User Interaction Recorder* (sequence of actions to get to the right contents, average time on each page to compute the attention applied by the visitor) finding existing patterns in data.

1) Algorithm used scanning the history of recorded user interactions.: This algorithm iterates two main steps.

The first step, called *harvesting user interaction statistics*, records the following list of measurements subdivided into two categories:

- tracks of the browsed page, like most requested pages, least requested pages, most accessed directory, average Time on Page, average Time on Site, ordered sequence of visited pages, new versus returning visitors (by means of cookies) and the number of views per each page.
- tracks of user activity at the page level:
  - Changing attributes of graphical elements: (e.g. changing charts representation from line chart to pie chart or histogram chart, sorting and filtering data in a table)
  - Removing/adding object elements (e.g. remove/add columns to the table)

The second step, called *grouping attributes of user interaction with significant correlation*, gathers in different subgroups the *tracking user activity* and *tracking browsed page* that have similar attributes, like most accessed directory and common graphics elements, in order to create mutually exclusive collections of user interactions sharing similar attributes.

To reach this goal, we use an algorithm handling mathematical and statistical calculations, such as probability and standard deviation, to uncover trends and correlations among the attributes of the user interaction.

For example, after scanning the history of recorded user interactions, an association rule "*the user that visits page one also visits page two and chooses to see histogram reports (90%)*" states that nine out of ten users that visit the page one also visit the page two and prefer to see the bar chart. We can buid use case models, based on these statistics, in order to reflect the requirements and the needs of each user. As a result, the user, classified under this use case, will take advantage to see a web interface based on his perspective, helping him to find and manage the information he needs more quickly.

2) Algorithm used to scan the PopCon logs.: PopCon is integrated within the *CMSSW* framework which depends on different tools like *POOL* and *CORAL*[4] and on database software like *ORACLE* and *SQLite*. This application can be used in two different ways:

- since it is integrated in the framework, users can write python scripts which are executed by the framework executable *cmsRun*.
- the framework itself provides an application which, using *PopCon* libraries, allows the exportation of data into the offline database.

These applications are responsible for maintaining and handling operations which are related to database transactions. In this scenario, it is very difficult to catch all error messages coming from different heterogeneous resources. Therefore, we follow this strategy: every application provides an error output consisting of three components: the name of application, the error code, that is unique for each tool, and the description of the error itself. So, PopCon developers can clearly understand what is wrong with their tool, while the end-user is able to check if the data exportation (database transaction) they want to perform was successful or not.

This error metric, for each tool, is provided by the framework developers in XML format in order to make it independent from the message sent to stdout and/or stderr.

Besides describing what the error is and how it occurred, most error messages provide advice about how to correct the error.

To help both users and developers to classify correctly the observed damage, the error messages are defined by the level of issue with a different colour. These levels are:

- Fatal. The program cannot continue (red colour).
- Major (Error). The program has suffered a loss of functionality, but it continues to run (orange colour).
- Minor (Warn). There is a malfunction that is a nuisance, but it does not interfere with the program's operation (deep green colour).
- Informational. Not an error, this is related information that may be useful for troubleshooting (green colour).

As further example, we describe another kind of error not depending on the particular application, but on Hardware/Software/Network problems. To discover this kind of error, we perform a time series analysis on database transactions associated with the discovery and use of patterns such as periodicity. Since dates and times of the database transactions are recorded along with the users information, the data can be easily aggregated into various forms equally spaced in time. For example, for a specific account the granularity of database transactions could be hourly and for other account could be daily. This information allows to discover two main kinds of alarm:

- Scanning the entities monitored by *PopCon* (logs of database transactions), the association rule "*during a long period, a specific user performs a database transaction at regular time intervals*" states that, probably, if these regular intervals suddenly change without a monitored interaction by an administrator, and, for particular cases, by the user, there can be network connectivity problems, or machine failures on the network. In details, if the system finds an exception to this pattern in data, it triggers an action to inform a user about possible problems by email. Besides, the web user interface provides red/orange/green alarms, according to the seriousness of the problem, so that this exception is immediately visible by the user.
- Taking the size of data together with the periodicity of database data transactions we can forecast the rate at which disk capacity is being filled in order to prevent a disk becoming full, alerting the database manager and the administrators of the machines dedicated to the data exportation some days in advance.

D. PopCon Info Collector

The *PopCon Info Collector* retrieves data from the *PopCon API DB* and the *PopCon User Interaction Recorder*. This component interacts with *PopCon Data-mining* to find existing patterns in data previously taken, and, finally, encodes them in JSON format, providing the result to the *PopCon Web Interface* (see figure 1).

E. PopCon Web Interface

The system has a front-end Apache server and back-end application servers. The PopCon Web Interface is an application created with a Python-based framework using `Cheetah`template engine to structure the web site. The *PopCon Web Interface* is built on the `CherryPy` framework application server, which runs behind Apache providing security module to automatically show a role-optimized view of the system and its controls. A set of reusable components, known as "widgets", are being made available. These are usually built using the `jQuery` libraries and are written in

[4]CORAL is a software toolkit (which is part of the LCG Persistency Framework) providing the set of software deliverables of the "Database Access and Distribution" work package of the POOL project.

CSS and JavaScript. Where possible, these are reused in order to provide identical functionality across direct components, so that a user feels comfortable with a standard style sheet for all web tools. The services run on a fairly standard configuration: a pair of Apache servers working as a load balanced proxy in front of many application servers. The front end servers are accessible to the outside world, while the back end machines are firewalled off from remote access[6]. With this infrastructure we can minimize problems related with security issues: in particular, each user is unable to handle database objects. Thanks to `AJAJ`[5] we can provide real-time feedback to our users exploiting server-side validation scripts, and eliminate the need for redundant page reload that is necessary when the pages change. In fact, this component allows to send requests asynchronously and load data from the server. The *PopCon Web Interface* uses a programming model with display and events. These events are user actions: they call functions associated to elements of the web page and then actions are recorded by the *PopCon user Interaction Recorder*. The contents of pages coming from different parts of the application are extracted from JSON files provided by the *PopCon Info Collector*.

IV. PopCon monitoring from the different users' perspectives

The design of the presentation of the data collected by *PopCon monitoring* is based on the requirements given by different types of users, each of them having to do with a different abstraction level of a Database administration issue: the ORACLE Database Administrator level, the central *CMS detector* level, the *CMS sub-detector* level and the End-User level.

- The ORACLE Database Administrator may wish to face up to databases security issues for which he is responsible. Typical example that can be detected:
  - people on the inside (using *PopCon tool*) and outside (using *PopCon Web Interface*) network who can access and what these users do;
  - programs accessing a database concurrently in order to avoid further multiple access to the same account;
  - if all such processing leave the database or data store in a consistent state;
  - illegal entries by hackers;
  - malicious activities such as stealing content of databases;
  - data corruption resulting from power loss or surge;
  - physical damage to equipment;
- The central CMS detector manager and the *PopCon tool* developer may require the possibility of analysing the behaviour of their applications for each CMS sub-detector.
- The sub-detector CMS manager may require the possibility to analyse the behaviour of his transactions on his own sub-detector database account.
- The End-User may require the possibility to analyse the behaviour of his own personal transaction such as size and rate/duration of the transactions, or detect fault situations related to insufficient password strength or inappropriate access to critical data such as metadata.

To summarize, *PopCon monitoring* automatically detects the cookies installed in each user's browser and this information is used to match the user with a role (Oracle Database Administrator, *PopCon tool* developer, sub-detector CMS manager, End-User) in order to provide a customized report that allows each user to have a customized printout of information depending on his needs.

The use of data mining techniques to extract patterns from logs of database transactions (operator, date and time) and the history of recorded user interactions has some general advantages. The storage of these patterns will help the user to read and understand quickly the current situation without going through several pages and use the search fields.

V. Conclusions

Although the number of samples analysed here is limited, the applied approach demonstrates that our open source application is dynamic since it can work and parse the different types of data for which date is a primary key.

Date can be written in many different ways because of flexible Python functions which work with date and parses it.

Another important feature of this application is that the *PopCon User Interaction Recorder* could be used in combination with *PopCon data-mining* to provide almost the same functionality in general for any application. It's indeed a flexible part which helps to collect and interpret information about user activities, and the actions made while he handles the application. This information can also be used to provide new and comfortable features for users, as we are using it to adapt the *PopCon Web Interface* to the user's needs.

References

[1] The CMS Collaboration. CMS Physics TDR, Volume I: Detector Performance and Software. Technical Report CERN-LHCC-2006-001; CMSTDR-008-1, CERN, Geneva, 2006.

[2] The CMS Collaboration. CMS Physics Technical Design Report, Volume II: Physics Performance. J. Phys. G, 34(6):995–1579, 2007.

[3] PopCon (Populator of Condition Objects). First experience in operating the population of the "condition database" for the CMS experiment. *International Conference on Computing in High Energy and Nuclear Physics, March 2009*

[4] CMS Computing TDR, CERN-LHCC-2005-023, http://cdsweb.cern.ch/record/838359 20 June 2005.

[5] The LHC Project. LHC Design Report, Volume I: the LHC Main Ring. Technical Report CERN-2004-003-V-1, CERN, Geneva, 2004.

[6] CMS conditions database web application service. *International Conference on Computing in High Energy and Nuclear Physics, March 2009*

[5] AJAJ: Asynchronous Javascript and JSON.

Using MPEG-21 to repurpose, distribute and protect news/NewsML information

Pierfrancesco Bellini, Ivan Bruno, Paolo Nesi
DISIT-DSI, Distributed Systems and Internet Technology Lab
Dipartimento di Sistemi e Informatica, Università degli Studi di Firenze, Firenze, Italy
http://www.disit.dsi.unifi.it/, nesi@dsi.unifi.it, paolo.nesi@unifi.it

Abstract

The distribution of news is a very articulated and diffuse practice. To this end one of the most diffuse formats for news production and distribution is the NewsML. The management of news has some peculiarities that could be satisfied by using MPEG-21 as container and related production tools and players. To this end, an analysis of modeling NewsML with MPEG-21 has been performed and reported in this table. The work has been performed for AXMEDIS project which is a large IST Research and Development Integrated Project of the European Commission.

1. Introduction

At present, there is a large number of content formats ranging from the simple files: documents, video, images, audio, multimedia, etc., to integrated content models for packaging such as MPEG-21 [1], [5], SCORM, MXF, NewsML [6], SportML, etc. These models try to wrap/refer digital resource/essences and in some cases to wrap them in a digital container, so as to make them ready and simpler for delivering. Among these formats the ones used for distributing and sharing news are mainly text and XML oriented such as NewsML of IPTC (International Press Tele-communication Council). Recently a new version of NewML has been proposed, the NewsML-G2 that provides support for referencing textual news, resource files, for paging them, while collecting metadata and descriptors, vocabularies, etc. (http://www.iptc.org). Furthermore, the news are typically massively processed by news agencies and/or by TV news redactions. They are not only received in NewsML formats but also in HTML, plain TXT, PDF formats as well. The agencies and redactions need to move, transcode, and adapt them to different formats processing both text and digital essences, by changing resolution, summarizing text, adapting descriptive metadata, etc. In some cases, the adaptation has to be performed on demand as a result of an answer to a query or request to a database or on a web service.

Moreover, frequently the news contain videos and images, while solution proposed by NewsML of zipping the file constrains the users to unzip the files in some directory to access and play the video. In addition, news contain frequently sensitive data for which protection of IPR (intellectual property rights) is needed. Thus, most of the above mentioned formats present a number of problems such as limitations related to the adopted packaging format. For example, such as the NewsML limitations on the packaging so as to prevent from playing effectively video content from the package without decompressing and/or unpacking, and limitations on the protection and preservation of the IPR (intellectual property rights). Such problems are related to the file format and protection support including certification, content signature and licensing.

Among the formats mentioned, the AXMEDIS implementation of the MPEG-21 file format and MXF supports the direct play. Only the MPEG-21 also support a range of business and transaction models via a DRM (Digital Rights Management) solution and with a set of technological protection supports.

In this paper, a solution to solve the above mentioned problems of news modeling, massive production and processing and distribution is presented. The solution proposed is based on AXMEDIS content model and processing GRID platform, AXCP. AXCP provides a set of technical solutions and tools to allow the automation of cross media content processing, production, packaging, protection and distribution. AXMEDIS multimedia processing can cope with a large number of formats including MPEG-21 and it can work with a multichannel architecture for the production of content on demand [3]. AXMEDIS is a framework that has been funded by the European Commission and it has been developed by many partners including: University of Florence, HP, EUTELSAT, TISCALI, EPFL, FHGIGD, BBC, AFI, University Pompeo Fabra, University of Leeds, STRATEGICA, EXITECH, XIM, University of Reading, etc. More technical information, as well as how to make registration and affiliation to AXMEDIS can be recovered on http://www.axmedis.org

In order to solve the above described problems, the AXCP solution has been augmented by semantic processing capabilities, NewsML modeling and conversion strategy into AXMEDIS MPEG-21 format with the aim of preserving semantics and capabilities of the early news files processed [4], [5]. In this case, the MPEG-21 models and tools have been used: (i) as a descriptor and/or a container (with AXMEDIS file format) of information and multiple file formats, (ii) as a vehicle to protect the IPR when the information is distributed towards non protected channels or it contains sensitive information.

The paper is organized as follows. In section 2, a short overview of AXMEDIS content processing platform for multimedia processing is reported. Section 3 refers to modeling of NewsML into MPEG-21 and AXMEDIS formats. In Section 4, some implementation details regarding the AXCP are reported. An analysis of the advantages identified in using the AXMEDIS model and tools are reported in Section 5. Conclusions are drawn in Section 6.

2. AXMEDIS Content Processing

The AXCP tool is based on a GRID infrastructure constituted of a Rule Scheduler and several Executors for process executing. AXCP Rules are formalized in AXCP java script [2], [4]. The AXCP Rule Scheduler performs the rule firing, discovering Executors and managing possible problems. The scheduler may receive commands (to invoke a specific rule with some parameters) and provide reporting information (e.g. notifications, exceptions, logs, etc…) to external workflow and tools by means of a WEB service.

The Rule Executor receives the Rules to be executed from the Scheduler and performs the initialization and the launch of the Rule. During the run, the Executor could send notifications, errors and output messages to the Scheduler. Furthermore, the Executor could invoke the execution of other Rules sending a specific request to the Scheduler, in order to divide a complex Rule/procedure into sub rules/procedure running in parallel, thus allowing a rational use of the computational resources accessible in the content factory, on the GRID. This solution maintains advantages of a unified solution and allows enhancing the capabilities and the scalability of the AXMEDIS Content Processing.

The AXCP processing tools are supported by a Plugin technology which allows each AXCP Rule Executor to link dynamically any content processing tool and algorithm (e.g. audio, video and image adaptation, transcoding, encryption) and to cope with possible customized algorithms and tools.

As to the processing capabilities, an AXCP Rule formalises in its own language features to perform activities of ingestion, query and retrieval, storage, adaptation, extraction of descriptors, transcoding, synchronisation, fingerprint, indexing, summarization, metadata manipulation and mapping via XSLT, packaging, protection and licensing in MPEG-21 and OMA, publication and distribution via traditional channels and P2P.

3. From NewsML to AXMEDIS modeling passing via MPEG-21

The NewsML has a structure at 4 nested levels (from the contained to the smaller components): *NewsEnvelope, NewsItem, NewsComponent* e *ContentItem (http://www.iptc.org)*.

The *News Component* mainly contains the information that may be used for modeling the *NewsItems*. At the end the *ContentItem* describes the contribution in terms of comments, classification, media type, format, notation, etc. The NewsML has also metadata mapped in the architecture and in particular in the *NewsComponent*: Administrative Metadata, Descriptive Metadata, and Rights Metadata. The information for the news identification are reported into the *NewsItems*, each of them can be univocally identified.

On the basis of our analysis, we have identified 6 main entities which have to be addressed: NewsML, NewsItem, NewsComponent, ContentItem, TopicSet, Catalog (see Figure 1)

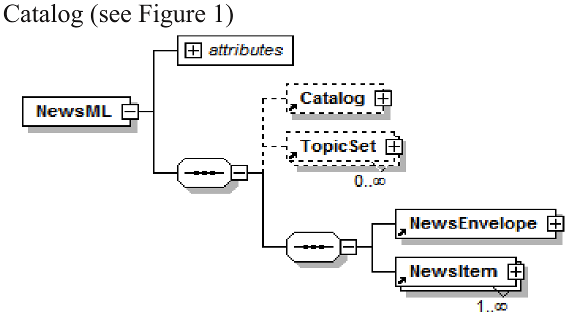


Figure 1 – NewsML main entities

The resulting model is hierarchical and in order to be ingested, analyzed and converted it has been replicated into an object oriented model allowing us to represent this model in the memory, by considering their relationships and roles, as in the UML diagram reported in Figure 2.

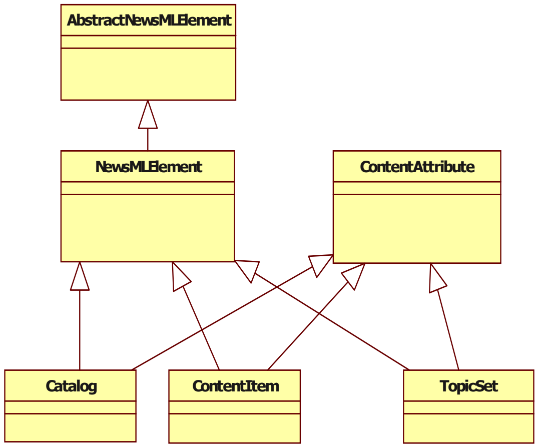


Figure 2 – Modeling NewsML main entities for conversion and analysis

In addition, also other classes have been implemented to model the NewsML such as: Topic, NewsMLDocument, NewsComponent, NewsItem also specialised from both NewsMLElements and ContentAttribute. The proposed model allows to ingest quickly the NewsML structures.

The realized model allows to perform the needed transformations on the NewsML files in an efficient manner. For example, the extraction of a NewsComponents removing its parts from the tree, the addition of news, etc, together with the conversions of the NewsML in other formats such as XML, HTML, Text and files, and MPEG-21 as described in the following.

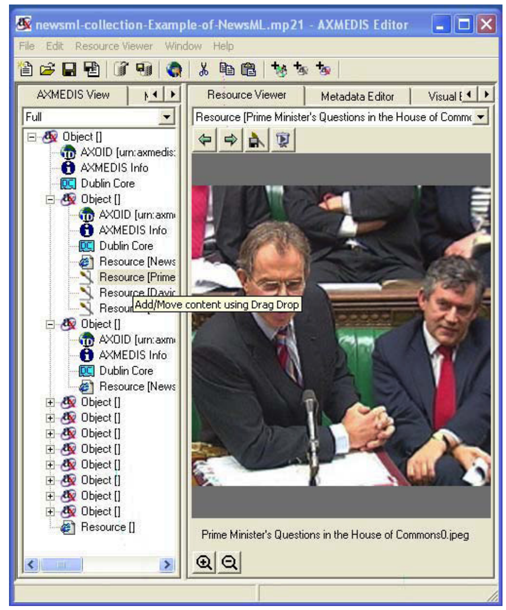
The resulted model has been also analyzed to map the information into the MPEG-21 structure of the DIDL (Digital Item Description Language).

NewsML element	AXMEDIS Element	MPEG-21
Metadata	AXInfo + Dublin Core	Descriptor
NewsML	AxObject	Item
NewsItem	AxObject	Item
NewsComponent	AxResource	Component
ContentItem	AxResource	Component

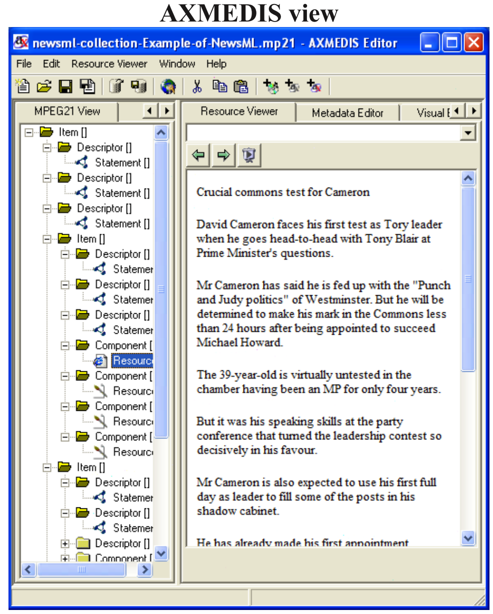
Table 1 – Mapping concepts of NewsML to AXMEDIS and MPEG-21

In Table 1, a mapping of the NewsML elements with those of MPEG-21 and AXMEDIS is provided.

The AXMEDIS editor allows you to see both MPEG-21 and AXMEDIS views of the newsML file as depicted in Figure 3. AXMEDIS view is only a more abstract view of the AXMEDIS file format ISOMEDIA based. The AXMEDIS mapping is more effective and easy to understand than the underlining MPEG-21 modeling that is fully flat and hard to be understood by humans. The resulted MPEG-21 container of the News can be protected by using the MPEG-21 REL and AXMEDIS tools for DRM.



AXMEDIS view



MPEG-21 view

Fig.3 – A NewsML on the AXMEDIS Editor

In Figure 3, AXMEDIS view, the nesting levels of AXMEDIS objects are evident. They can be moved or extracted simply using drag and drop. The same approach can be adopted to work with single contributions: text and/or digital files (images, video, etc.). They can be played directly into the editor and into the AXMEDIS player. An additional feature is the index in HTML of the converted NewsML items. It has been automatically produced by processing the NewsML structure in the AXCP script. That index is an HTML file enforced into the AXMEDIS Object (see the bottom of the tree in Figure 4).

4. Implementation on the AXCP GRID

The above mentioned object oriented module for NewsML ingestion, modelling and processing has been added to the AXCP Node engine. Therefore, a set of functionalities, API, to access the NewsML models has been defined and made directly accessible into the AXCP Java Script Multimedia processing language.

5. Benefits and results

This solution based on AXCP allowed to set up flexible automatic processes where NewsML information is ingested and processed in a very efficient manner, while considering any kind of conditions and structures for repurposing them and adapting news including text and digital essences towards different formats: HTML, TXT, PDF, MPEG-21, SMIL, etc., either integrating or not digital essences into them and distributing them via email, posting on FTP, on DBs, etc.

Besides, the news modeling with AXMEDIS has some advantages, as the resulting AXMEDIS object can be:

- used as a news descriptor and/or a news container (with AXMEDIS file format), supporting any kind of file formats for the digital essences being integrated into the news.
- used to manipulate the news, to add other information via AXMEDIS Editor and to make a directly play of the essences into the news without extracting them from the package.
- searched into the internal body of the news object, thus making the understanding and browsing of complex news easier, by adding simple Intelligent methods such as the ones described into [5]. .
- annotated conformant to MPEG-21 as described in [5].
- IPR protected when the information is distributed towards non protected channels or it contains sensitive information.
- distributed in several manners and accessed via PC, PDA, etc.

6. Conclusions

In this paper, the analysis of the modelling NewsML and news in genral with MPEG-21 has been performed and presented. The results demonstrated that the structure of the News can be quite easily modelled in MPEG-21. In addition, the news processing consisting in their ingestion and transcoding can be performed on the AXCP platform in quite easy manner since now an ingestion module of NewsML has been developed and added. As a result, a number of advantages have been identified and demonstrated, as reported in Section 5. The full documentation can be recovered on the AXMEDIS portal http://www.axmedis.org. AXMEDIS is an open platform, which means that you can join the AXMEDIS community. The example mentioned in this paper is accessible from the same web portal.

Acknowledgments

The authors would like to express their thanks to all the AXMEDIS project partners including the Expert User Group and all the affiliated members, for their contribution, funding and collaboration efforts. A specific acknowledgment to EC IST for partially funding the AXMEDIS project. A warm thanks to all the AXMEDIS people who have helped us in starting up the project and sorry if they might have been not involved in the paper nor mentioned. We trust in their understanding.

References

[1] MPEG Group, "Introducing MPEG-21 DID", http://www.chiariglione.org/mpeg/technologies/mp21-did/index.htm

[2] J. Thiele, "Embedding SpiderMonkey - best practice" http://egachine.berlios.de/embedding-sm-best-practice/embedding-sm-best-practice-index.html

[3] P. Bellini, I. Bruno, P. Nesi, ``A language and architecture for automating multimedia content production on grid", Proc. of the IEEE International Conference on Multimedia & Expo (ICME 2006), IEEE Press, Toronto, Canada, 9-12 July, 2006.

[4] P. Bellini, P. Nesi, D. Rogai, "Exploiting MPEG-21 File Format for cross media content", Proc. Of the International Conference on Distributed Multimedia Systems, DMS 2007, September 6-8, 2007, San Francisco Bay, USA, Organised by Knowledge Systems Institute.

[5] P. Bellini, I. Bruno, P. Nesi, M. Spighi, "Intelligent Content Model based on MPEG-21", in Proc. AXMEDIS 2008, Florence, Italy, 17-19 Nov. 2008, pp 41-48, IEEE Press.

[6] Kodama, M.; Ozono, T.; Shintani, T.; Aosaki, Y.; "Realizing a News Value Markup Language for News Management Systems Using NewsML" Complex, Intelligent and Software Intensive Systems, 2008. CISIS 2008. International Conference on 4-7 March 2008 Page(s):249 - 255

Activity-oriented Web Page Retrieval by Reflecting Human Traffic in the Real World

Atsuo Yoshitaka*, Noriyoshi Kanki**, and Tsukasa Hirashima**

*School of Information Science
Japan Advanced Institute of Science and Technology
1-1 Asahidai, Nomi, Ishikawa, 923-1292 Japan

**Graduate School of Engineering
Hiroshima University
1-4-1 Kagamiyama, Higashi-Hiroshima, Hiroshima, 739-8527 Japan

Abstract

Currently, major sources of information are not only in the real world but also in the information space organized by WWW on the Internet. Information acquisition and retrieval related to the real world need to recognize user's behavior in order to fill his/her needs. In this paper we present behavior oriented information retrieval system and its experimental operation. Users' activity in the real world, i.e., trajectory projected onto geographical map with indices of places, is tracked by GPS receivers. Commonly and frequently observed movements by users are detected and they are applied in the process of evaluating the importance of information to be retrieved that relates to places or facilities in the real world. The proposed system assists a user to behave in the real world in the sense of retrieving information that helps to decide his/her subsequent actions to take.

1. Introduction

Mobile computers which are small enough with high computing capacity became widely available. The diffusion of these devices is one of the dominant factors that support recent mobile computing environments. In recent years, researchers are studying real world oriented information management especially in mobile environment. This kind of information management ranges from information acquisition to information provision. One of the directions is to provide a user with information that is related to his/her current activity in a certain time, place and/or occasion. Sensing the context of a user's activity is achieved by tracking the users movement projected onto a geographical map. Various sensors are available to capture this kind of activities. GPS (Global Positioning System) receiver is widely used in order to track the activity of a user. If it is possible to assume that objects being accessed or manipulated are limited and modified in advance, RFIDs may be implanted into the object for tracking user's behavior or movement.

For the advanced real world oriented information management, we believe the system should be capable of managing information in accordance with user's context of activity. In the process of information management in the real world, one of the most important features is the method of recognizing the target of the user's scope of interest for information indexing, filtering or retrieval[1]. With respect to information provision, the sources of information is often a dedicated information storage obtained by the process of information acquisition.

However, we should aware that the range, quantity and sometimes quality of information accessible via WWW is not negligible, and various real-world related information is provided by individuals, shops, companies and so on. However, most of the existing web retrieval interfaces are not taking the context of human movement in the real world into account. Related to this issue, some web interfaces are proposed that project web pages onto geographical map and let a user to access various web pages related to shops, train stations, buildings, event halls and so on, associated with icons on a map presented on a mobile computer display. However, as far as we know, there are no work that reflects users' context in the sense of activities in the real world.

In this paper, we describe a framework that accumulates users' activity corresponding to places or facilities where he/she stayed with a certain purpose, and retrieves information related to the situation he/she is facing from the Web. In this study, we regard WWW as the public information storage and propose a framework of context aware Web retrieval based on users' activity from the point of view of traveling from one place to another. Web contents are retrieved based on the accumulation of activities of either a group of users or a user to be assisted.

From the point of view of accessing Web information, non-context aware retrieval where the target of retrieval does not reflect the user's current location or movement may require a number of trials of refinement of specifying keywords to submit to a search engine. On the contrary, proposed framework, i.e., context aware retrieval implicitly provides the search engine with additional keywords that represents the user's expected destination of movement as well as current location in the real world.

2. Behavior Modeling in the Real World

2.1 User Activity Model in Mobile Environment

Recently, most of the companies, shops, public places such as city libraries, concert halls, train stations, and so on, provide information related to themself on the Web. In addition, portal sites on shopping, travel, cuisine, entertainment and personal blogs are also nonnegligible source of information describing such facilities. They often update their Web pages and provide us up-to-date information, and the Web contents often provide us information that may affect decision of our activity in the real world.

From above-mentioned point of view, we assume the user activity model as follows in mobile environment in this paper.

(1) A user moves from one place or facility to another in accordance with a certain reason such as business, travel or pleasure.
(2) During the activity, he/she retrieves information on the candidate place or facility where he/she is going to visit subsequently. In this process less steps of manipulation or keywords are preferable for easiness of use.
(3) The user accesses the information on the candidate place to visit subsequently, and makes a decision to visit there or not, or changes the destination to visit by referring to the information.

The idea of user activity model is illustrated in Figure 1. This user activity model may be regarded as a general situation in information retrieval on facilities or places to decide one's behavior in mobile environment. In the following sections, we concentrate on information retrieval following above-introduced activity model.

In modeling user's activity, we focus on origin-destination oriented movement, i.e., movement from one place (or facility) to another, regardless of the route between them. This is based on the assumption that his/her subsequent action of movement is affected by his/her current location. Attributes on the facilities or places in the real world, such as name, postal address, or the type of service, correspond to 'keywords' for information retrieval in the case of the above-mentioned scenario. We regard facilities where more traffic of users exists between one's current position to them are the 'near' places he/she may

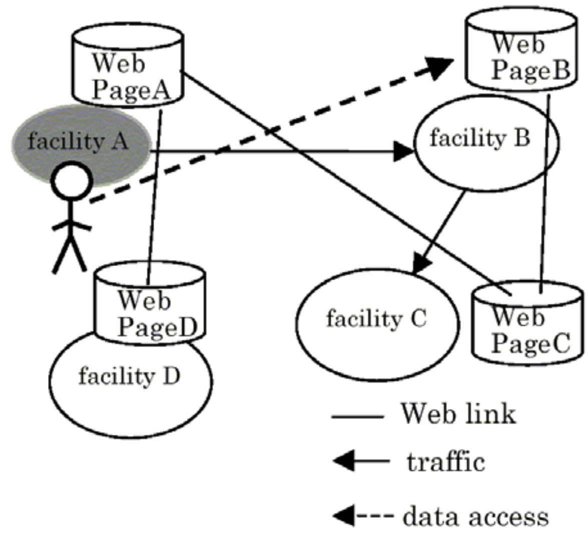


Figure 1. Relation between movement and retrieval

visit as the next action, the information related to them is more important than others for the decision of the next destination. Note that this idea is not based on geographical distance but logical distance based on the frequency of human traffic between one place to another. It means if more traffic between place *A* to *B* is observed than that between *A* to *C*, where *B* is further than *C* from *A*, we regard the information related to *B* is more important than that related to *C* for a user whose current position in the real world is *A*. This idea is different from the idea of geographical distance based information filtering or retrieval.

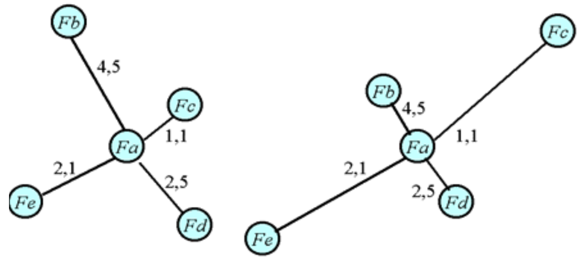
In the subsequent sections, we describe the framework of information filtering based on users activity in the real world.

2.2 Traffic Graph

We assume a user's task of retrieving Web pages as part of the activity in the real world. In this context, the user's objective of retrieving Web pages is to obtain information that is related to a facility such as a store, a train station, a school, a city hall, and so on. It is popular that various facilities provide public with information on timely events or notice via Web pages. This kind of information is valuable in the sense of deciding his/her subsequent action.

Based on this observation, we discuss detecting the facilities where a user stayed for a certain purpose. Detail of the criteria on detection of staying is described later in this paper. Based on the detection of the facility where a user stayed, we extract a user's stay at a place to model the traffic between facilities for context aware Web retrieval.

The basic idea is based on the observation as follows. Assume that a person is currently staying at a place



(a) location based arrangement (b) traffic based arrangement

Figure 2 Traffic Graph

associated with a place (i.e., facility) *Fa* and trying to access Web pages in order to obtain information for the place where he/she is going to visit subsequently. Under the assumption where frequent traffic, i.e., users' movement from one place to another, is observed between *Fa* and *Fb*, we extrapolate that he is going to retrieve Web pages related to the facility *Fb*.

In order to model the users' traffic between facilities, we introduce *traffic graph*. An example of traffic graph is illustrated in Figure 2. In the figure, a node denotes a facility (i.e., a place) in the real world. The first element in the pair of values attached to a link represents geographic distance between facilities and the second one represents traffic frequency. In figure 2(a), the length of a link corresponds to geographic distance between facilities in the real world. On the other hand, the length of a link corresponds to the closeness with regard to traffic between two places to stay in figure 2(b); the more the traffic between the facilities, the closer they are in the sense of travel frequency.

Here we regard higher travel frequency corresponds to higher possibility of needs of information related to the facility where he/she is going to visit. In the above example, we regard Web pages related to *Fb* and *Fd* are more expected to be accessed than *Fc* and *Fe*, under the assumption where the current location of a user demanding Web information is *Fa*. The traffic graph is organized by tracking the movement of multiple users. That is, the history of traffic is shared by multiple users in order to derive traffic density between facilities for evaluating importance of Web information in the sense of the human traffic-based relation between facilities. Based on the traffic graph we measure the importance of Web pages that corresponds to the context of users' activity. In the process of organizing a traffic graph, users to share traffic history may be grouped based on the preference of individual, and the group is dynamically reorganized in accordance with the transition of activity context. Privacy issue can be avoided by anonymizing individual traffic data.

3. Extraction of Human Behavior in the Real World

3.1 Activity Tracking

A user's position in the real world is traced based on positioning data from GPS (Global Positioning System) receiver. GPS system detects the current position by evaluating the temporal difference of radio wave received from several satellites; more radio waves are received, more precise position is detected. That is, the error distance between the detected position and the true location where a GPS receiver is placed varies depending on the radio wave condition. Since data from a GPS receiver consists of coordinates by longitude and latitude, the coordinates data is projected onto geographical map with latitude-longitude index and rectangular regions corresponding to facilities such as schools, shops, restaurants, public halls, and so on. Each rectangular region is associated with the description that consists of textual description of address and the name of the facilities.

In order to organize a traffic graph, we need to extract a place when a user came over. As stated, detected position by GPS contains the error of distance, whose amount depends on the condition on receiving satellite waves. Therefore, this error needs to be taken into account to diminish misdetection. The positioning error is generally estimated as 2*drms*, where *drms* stands for distance root mean square. The error in the positional data e_p is estimated by the following formula.

$$e_p = 2drms = 2UERE \times HDOP \quad (1)$$

In the above formula, *UERE* is the abbreviation of user equivalent range error, which is not obtained from GPS data. This value is determined as 2.0 assuming general open-air condition. *HDOP* stands for *horizontal dilution of precision*, which is obtained from GPS data. The value of *HDOP* approximately ranges from 1 to 2 where the receiver can get enough number of satellite waves. In case where the wave condition is not satisfactory, it ranges from 7 to 9. Therefore, average positioning error in good wave condition is approximately 6 meters, and it is approximately 36 meters in bad condition. The user's activity in the real world is detected not based on trajectory obtained by GPS data, but on the places where he/she stayed, following the user activity model described in 2.1.

3.2 Detection of Stay

As discussed in 3.1, we consider the place where a user stayed for more than a certain duration is a distinctive place to analyze his/her activity. Our objective is to detect mutual strength between places in the sense of human traffic. Therefore, not tracing user's activity by means of GPS coordinates themselves but detection of whether he/she stayed at a place need to be obtained.

The state of staying at the place of a facility is detected by taking the size of facility as well as the positioning error into account. Based on the pre-experiment, we determined the threshold $t_{stay}(p)$ with regard to the facility p for classifying whether a user stayed at a place or not.

$$t_{stay}(p) = \frac{L_R + 2e_p}{v_{walk}} \quad (2)$$

In the above formula, L_R denotes the sum of length and width of a minimum rectangle that covers the area of a facility. The walking speed of a user is denoted as v_{walk}.

The judgment of stay is carried out as follows. First, when the user is located at a position whose distance from the nearest edge of the rectangular region of a facility is less than e_p, the duration of stay is started to be measured until he/she moves away from the region. If the duration exceeds $t_{stay}(p)$, his/her activity is classified into 'stayed' at the facility p.

We experimented to evaluate the performance of detection of user stay based on precision and recall. We assumed that all the places of facilities in the route of user traffic are defined as the part of geographic data in advance. Let N_{stay} denote the number of actual staying at places that a user made in their activity. We denote the number of extracted stay at facilities by this method and that of correct stay at places within the detected stay as $N_{extracted}$ and $N_{correct}$, respectively. The number of actual stay at places is denoted as N_{actual}. The precision and recall of detecting stay at a facility, $Precision_{stay}$ and $Recall_{stay}$ is represented as follows.

$$Precision_{stay} = N_{correct} / N_{detected} \quad (3)$$

$$Recall_{stay} = N_{correct} / N_{actual} \quad (4)$$

According to the result of 30 days of experiment to track a user's activity, where the user was a graduate student, the system detected 132 times of stay at facilities. When the positioning error of GPS is not taken into account in the process of detection, precision and recall were 0.92 and 0.67, respectively. In case where the positioning error is taken into account as we described, the precision and recall were 0.88 and 0.76, respectively. As a consequence, stay detection with GPS error adaptation improves recall with little degradation of precision. This performance may be improved further by taking the direction of motion trajectory into account.

4. Situation Aware Web Retrieval

Context aware Web retrieval is performed based on the Traffic Graph. As described, geographical map data is prepared with the region of a facility, and each region is associated with the name and the postal address of the facility. It is popular that a Web page related to a facility contains the description of the postal address as well as the telephone number for the help of visiting. Therefore, empirically speaking, the possibility of desired Web pages being listed at the highest position of the ranking in the result of retrieval increases by appending the keywords of the name of facility and the address of it. Figure 3 shows

Figure 3 Overview of the Prototype System

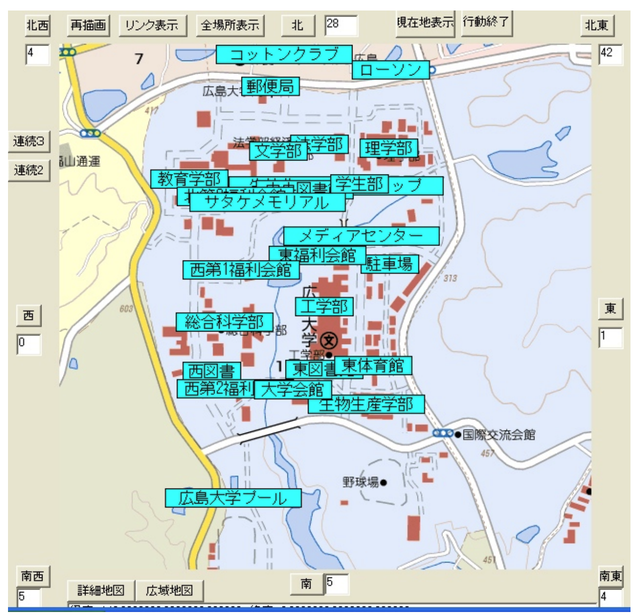


Figure 4 Displaying Nodes

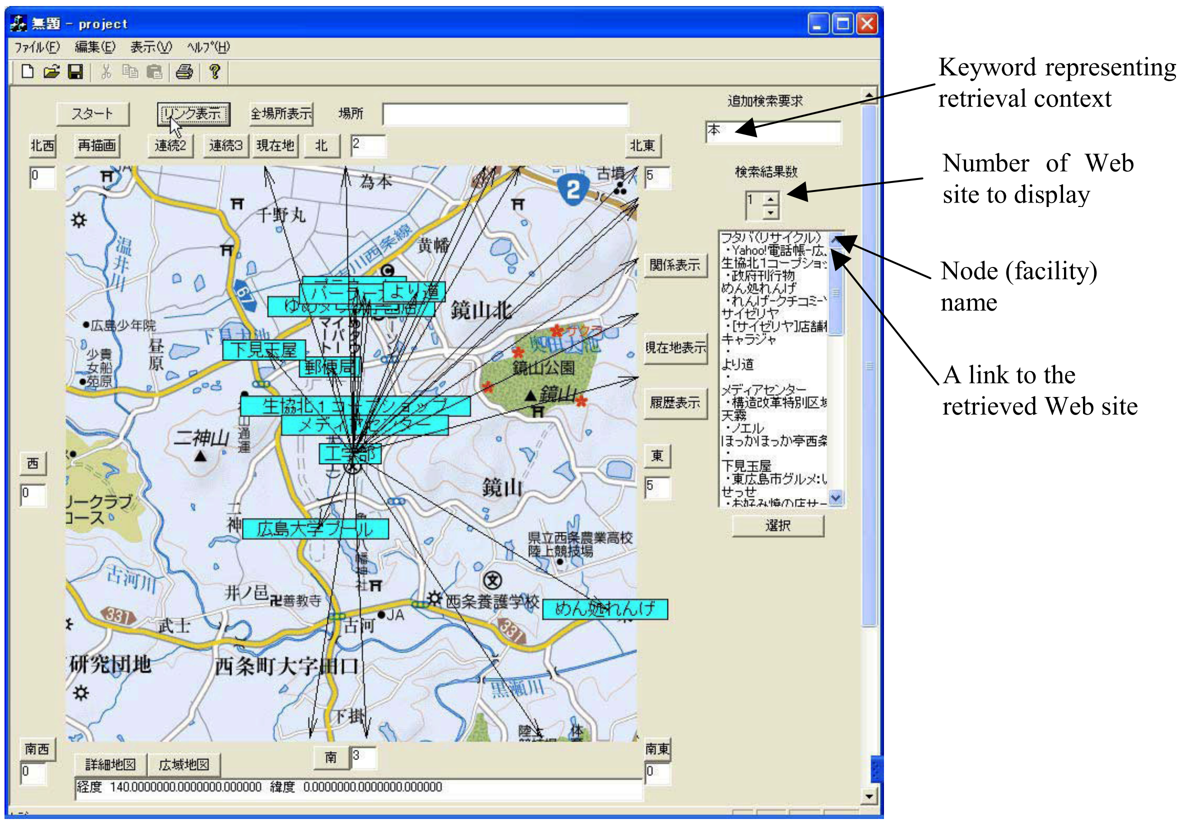


Figure 5 Displaying a Traffic Graph

the overview of the user interface of the situated aware web browser. The system is implemented on the SONY VGN-U71P with Visual C++. The small black module above the PC is a GPS receiver, which is connected to the PC via USB. Manipulation by a user is performed with a stylus pen. Figure 4 shows all the nodes near the current location, which correspond to facilities or places, registered in the system. The name of the facility is denoted in each of the nodes in Japanese letter. The scale of the map may be changed if needed.

Figure 5 is an example of showing detected traffic between facilities, each of the traffic is shown as a link between facilities. The name of a facility, such as the name of a store, a school, and so on, is displayed in a rectangle. In this example, the current position of a user is displayed at the center of the map.

When a user enters a keyword in the upper right text box in Figure 5 for Web page retrieval (in the figure, “本”, i.e., book, is specified), the name and the address of the facilities where the traffic from/to the current position exist are appended to the keywords that the user entered explicitly. This process is repeated for all the facilities where traffic to/from the current position of facility is detected. The top *n* facilities (or places) in the result of the retrieval for each of traffic are arranged in the descendant order of the frequency of traffic, and they are presented to the user. The result of the retrieval is displayed in the form of list of the titles of Web page at the right of the interface in Figure 5. Simply clicking one of the title in the list, the Web page corresponding to it is displayed. Figure 6 shows the difference between an ordinal retrieval by specifying a keyword for the Google on the right and that of the proposed method. Though they are displayed in Japanese, Google search with a single contextual keyword such as ‘noodle shop’ returns well known portal sites that describe noodle restaurants all over Japan, bulletin board sites on noodle, or a link to the article of noodle in Wikipedia. On the other hand, the retrieval result by the proposed method only shows Web pages related to noodle restaurants where many local inhabitants often visit, which are denoted with underlines.

In the implementation of the functionality of context aware Web page retrieval, we post HTTP request to the Google search engine. When a user tries to retrieve Web pages by specifying keywords, his/her current position is detected by matching GPS data and map data, and the name and the address of the facility where he/she is staying is extracted. The HTTP request to the Google search engine is invoked by posting extracted name and address of the facility as well as explicitly specified keywords. This framework enables a user to retrieve Web pages related to the facilities where frequent travel from the current location is observed. This, in turn, provides context aware Web retrieval reflecting human activity in the real world, based on the idea that frequent travel between places implies higher priority or importance of information to be retrieved, in case of searching for information to decide the subsequent actions.

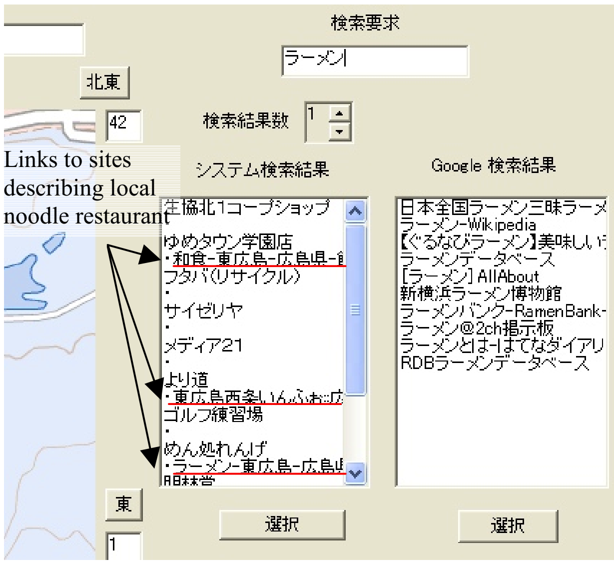


Figure 6 Comparing the Result of Retrieval

4. Related Work

The objective of context-aware web browser is to adapt the variety of needs or purposes of the users [2-4] in information retrieval. However, the direction of these studies is different from one to pursue adapting information provision with regard to a user’s activity in the real world, i.e., the activity in the real world is not taken into consideration as criteria in information retrieval.

Situation aware, i.e., location dependent, Web browsing is studied in [5] and [6]. In [6], GPS signal is referred to for acquiring position in the real world for location dependent Web browsing. However, it does not reflect the traffic or flow of persons in the real world in evaluating the importance of Web information related to the human traffic. Therefore, we classify it as static, location oriented Web browsing, which does not take dynamic human traffic into account. As far as we know, there is no study that retrieves Web pages based on human traffic between places, i.e., the context of user activity in the real world.

5. Conclusion

We described a novel framework of information retrieval based on dynamic user's activity in the real world. In the framework, the density of human traffic from one place to another is regarded as the strength of relation in the sense of information as well. However, this method of valuation may be biased with reference to the size of city, public traffic network, which is one of the open issues to be investigated.

When this system is operated in a large scale, traffic history of each user will be accumulated in a mobile computer and it is transmitted to a central sever via wireless network in order to construct traffic graph. We employed GPS receiver to detect the place or facility where a user stayed. Currently, this method is a reasonable option that can be widely operated. However, it cannot detect exact destination in case of complex indoor field or buildings. It might be replaced with another method such as utilizing widely diffused RF-tags in the future.

Proposed framework enables to retrieve desired information with less keyword to specify for accessing search engine. Additional, situated keywords are implicitly applied in addition to the explicit keyword given by the user, which improves the hit-ratio and diminishes the cost for accessing requisite Web pages.

Acknowledgement

This work is partially supported by the Grand-in-Aid in Scientific Research, JSPS.

References

[1] P. J. Brown and G. J. F. Jones, “Context-aware Retrieval: Exploring a New Environment for Information Retrieval and Information Filtering,” Personal and Ubiquitous Computing, Vol. 5, Issue 4, pp. 253-263, 2001.

[2] G. N. Prezerakos, N. D. Tselikas, G. Cortese, “Model-driven Composition of Context-aware Web Services Using ContextUML and Aspects”, Proc., IEEE International Conference on Web Services, pp. 320-329, 2007.

[3] A. Thawani, S. Gopalan, and V. Sridhar, “Web-based Context Aware Information Retrieval in Contact Centers”, Proc. International Conference on Web Intelligence, pp. 473-476, 2004.

[4] T. Koskela, N. Kostamo, O. Kassinen, J. Ohtonen, and M. Ylianttila, "Towards Context-Aware Mobile Web 2.0 Service Architecture", Proc., International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, pp 41-48, 2007.

[5] A. Haghighat, C. Lopes, T. Givargis, and A. Mandal, "Location-Aware Web System", Proc., Workshop on Building Software for Pervasive Computing, OOPSLA'04, 2004.

[6] D. Carboni, S. Giroux, et al., "The Web around the Corner: Augmenting the Browser with GPS", Proc., the 13th International WWW conference, pp. 318-319. 2004.

An Architecture for User-Centric Identity, Profiling and Reputation Services

Gennaro Costagliola, Rosario Esposito, Vittorio Fuccella, Francesco Gioviale
Department of Mathematics and Informatics
University of Salerno
{gencos,vfuccella,cescogio}@unisa.it, scalax.uep@virgilio.it

Abstract

This paper presents a work in progress whose objective is the definition of a novel architecture for solving several challenges related to Web navigation, such as accessing to multiple Web sites through a single identity and verifying the identity and the reputation of a peer involved in a transaction. The proposed model tries to solve the above challenges in an integrated way through the introduction of a specialized Web Mediator acting on behalf of the user during usage of the Net, identity providers for identity data centralization, and a two way negotiation system among parties for mutual trust.

1. Introduction

The need for introducing new functionalities to improve the user Web experience is more and more widely felt. Lately, researchers are closely taking into account the following important issues:

1. Registering and accessing to multiple services using a single identity for all services (single sign-on systems);

2. Verifying the identity and the reputation of a peer (user or organization) involved in a transaction;

3. Keeping the property and control of personal information such as: user profile, reputation, etc;

In this paper we propose an architectural model aimed at pursuing the above objectives trough the introduction of a *Web Mediator* (*WM*) acting on behalf of the user during Web navigation and an *Identity Provider* for the identity data centralization. The former is responsible for maintaining user personal data and profile to use in content personalization (as similarly done in [1]). The latter is responsible for keeping user identity and reputation data, and to vouch for the user in registration and authentication procedures.

Our model enables a two way negotiation system among parties for mutual trust: in a transaction both parties can mutually authenticate and verify reputation and profile. This sort of handshake, will allow them to decide whether the transaction can go on or should stop. It is worth noting that despite adding new functionalities to the actual Web application interactions, the architecture works with the actual Web protocols.

The advantages deriving from the availability of a solution to the three issues mentioned before are evident in several scenarios occurring daily during Web navigation. For instance, mutual trust is useful in the detection of phishing: let us suppose a user receives an e-mail containing a link to an important document about his/her bank account stored on the bank Web site. By connecting to the link with our framework enabled, the user can both check whether the remote Web server supports the architecture and verify its credentials. The phishing attempt can be immediately detected in the former case and after a reputation check in the latter case. The availability of user profile and reputation is useful in many cases: i.e., profile is used for offering personalized services, reputation in on-line auction services. Their availability to the user is advantageous since: data are already available when a user starts requesting a service at a new provider (it is not necessary to wait for a new profile or reputation to be built); the user is owner of his/her personal data which can be used with different sites offering the same services.

The above mentioned issues have been faced separately so far, that is, to our knowledge, there are no proposals of a generic architecture offering a solution for them all in literature. I.e., platforms for single sign-on [6] trust and reputation management [3] are available, as well as methods for preventing phishing [5]. In order to propose a unique solution to the above challenges, we have decided to extend a well established *SSO* platform, *OpenID* [6], with the support of a mutual trust establishment procedure. In particular, we have extended the *OpenID* Authentication procedure. The interaction among user's and peer's modules involved in the procedure are described through the paper.

In our prototype, the Web browser can communicate with user's *WM* through a special plug-in.

The rest of the paper is organized as follows: in section 2, we introduce the *OpenID* platform; the architectural model, including a detailed description of the involved entities and their interaction model, are presented in section 3. In section 4, we will describe the implemented prototype and its instantiation in a real-life application scenario. Final remarks and a discussion on future work conclude the paper.

2. The OpenID Platform

OpenID was firstly developed in 2005 as a user-centric and URI-based identity system. Its main objective was to support the *SSO* functionality. The initial project has grown and has evolved in a framework enabling the support of several functionalities which can be added to the basic platform.

The *OpenID* architecture components are: the user, the remote Web-server (also know as *Relying Party*) where the user wants to authenticate and the *Identity Provider* (*IdP*) that provides vouch for user identity certification. *OpenID* has a layered architecture. The lower layer is the *Identifier* layer. This layer provides an unique identifier for address based identity system. The address identifier (*OpenID* URL) is used by the *Relying Party* (*RP*) to contact the user's *Identity Provider* and retrieve identities data. Both URL and XRI [7] address formats are supported as identifiers.

The above layer is the service discovery layer. It is implemented trough the *Yadis* protocol [4]. The purpose of this layer is to discover various type of services reachable through an identifier. In the case of *OpenID* it is used to discover the *Identity Provider* location.

The third layer is the *OpenID Authentication*. The main purpose of this layer is to prove that an user is the owner of an *OpenID* URL and, consequently, of the connected user data.

The fourth layer is the *Data Transfer Protocol*. This protocol is used to transmit user related data from the *IdP* to the *RP*. In *OpenID Authentication 1.1* this layer is implemented through the *SREG* protocol (*Simple Registration Protocol*), which allows the transmission of simple account related data [2]. Currently, the *OpenID* research community is defining a new version of the protocol capable to transmit various type of data other than identities related one.

3. The architecture

In this section we give a description of the proposed architectural model, including the involved entities and their interactions in a trusted negotiation, which is a typical interaction where two parties gradually establish trust [8]. It

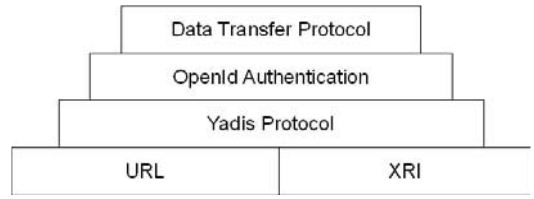


Figure 1. The *OpenID* layered architecture.

is based on the previously described *OpenID* platform, and extends it to support the features outlined in the introduction.

Our model extends the *OpenID* platform by enabling the establishing of mutual trust and the exchange of reputation and profile data between two parties. In particular, it adds *Profile* and *Reputation* layers upon the uppermost *OpenID* layers and a *Mutual Trust* layer above them (Fig 2).

Reputation management service is provided as an extension of the *DTP* layer. In particular, the data model supported in the information exchange occurring at this layer is extended with reputation data. The discussion on how to represent, create and manage these data are out of the socpe of this paper and will not be treated here.

User profile data are managed by the *WM*, which also works as a profile provider, and can be accessed only after the *OpenID* Authentication procedure is successfully completed.

The *Mutual Trust* layer implements the handshake procedure that will authorize the user application to proceed with an interaction after identity, reputation and profile of remote peer are checked.

In a typical scenario, our architecture is composed of the following components:

A) The Web Browser equipped with a specific plug-in (i.e. a Firefox add-on) to communicate with the *WM*;

B) A *Web Mediator* (*WM*): the software module responsible to communicate with other remote peer *WMs*, in order to perform a trusted negotiation. The *WM* can perform two functions: issue a transaction request to remote peers *WMs* or receive incoming transaction requests from remote peer *WMs*. In the case it is the first to send a request will refer to the *WM* as *User Web Mediator* (*UWM*); otherwise we will refer to it as *Remote WEB Mediator* (*RWM*). More in details, a *WM*, by referring to a preference table set by the user, verifies the identity, reputation and profile of remote peers and, after that all checks are passed, it authorizes the application to proceed with the transaction. Furthermore, in scenarios that needs this feature, it also checks that the resource retrieved as a transaction result fits user's preferences (i.e. content filters).

C) An *Identity Provider* (*IdP*), deployed on a third party server, that is responsible for guaranteeing the veracity of the credentials issued by the *WMs*; it is also responsible to

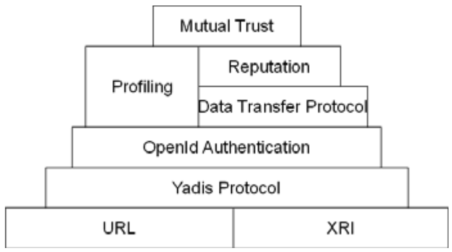


Figure 2. The proposed architecture.

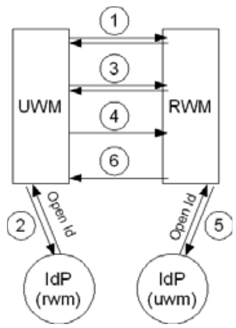


Figure 3. The WM Handshake

provide, by extending the common data already passed during an *OpenID* authentication, the reputation data.

D) The remote application that provides the requested resource after being authorized to do so from the *RWM*.

Before we start to discuss the fundamental phases that occurs in a transaction we will describe the *WM Handshake* procedure between *WMs* in which *UWM* and *RWM* proceed to establish a mutual trust with the help of one or more *IdPs*. During this phase the *WMs* exchange profile and reputation data and verify that the user parameters are satisfied. More in details, as shown in figure 3:

1. *UWM* requests the *OpenID* URL to the RWM and receives it;

2. *UWM* starts the authentication procedure by contacting *RWM*'s IdP which authenticates *RWM* and replies with the *RWM*'s reputation data;

3. *UWM* recovers *RWM*'s profile data trough a GET request to the *RWM* using a standard URL;

4. *UWM* checks the received profile and the reputation data and, if all checks are passed, sends its *OpenID* URL to *RWM*;

5. *RWM* starts the authentication procedure by contacting *UWM*'s IdP which authenticates *UWM* and replies with *UWM*'s reputation data;

6. *RWM* recovers *UWM*'s profile data trough a GET request to the *RWM* using a standard URL;

7. *RWM* checks the received profile and the reputation data and, if all checks are passed, sends an OK message to *UWM*.

The authentications in step 2 and 5 follow the *OpenID* protocol and consist of sending username and password to the *IdP* (through a POST request) to prove to be the owner of the identity related to the previously sent *OpenID* URL.

For sake of clarity no exceptions are shown in the procedure. In the case something goes wrong, the *UWM* is the one in charge of notifying the user application that the handshake did not succeed.

Note that, by following the previous steps, *UWM* is the first to see the other's reputation and profile data. Furthermore, the *RWM* will be able to access to the *UWM* data only if it is considered worth to receive it. This is the *UWM-first* version of our architecture. The *RWM-first* version is easily obtained by letting the *UWM* start sending its own *OpenID* URL and modifying the next steps accordingly.

In the following, we describe the complete transaction between two Web applications (user and remote applications) by following the *UWM-first* approach (the other case can be easily derived). More in detail, as shown in figure 4:

1. the user makes a request to the application to execute a transaction with a remote application;

2. the user application contacts its *UWM* to obtain an authorization for the transaction;

3. the *WM* Handshake between the corresponding *UWM*, *RWM* and *IdPs* occurs as described above;

4. if the handshake succeeds, the *UWM* sends the shared *RWM OpenID* authorization token to the user application;

5. the user application sends its original request together with the authorization token to the remote application;

6. the remote application uses the token to query its RWM for the identification and profile of the requester (as built with the *UWM*);

7. the *RWM* returns the required resource; from now on the transaction between the two applications does not involve the underlying levels.

In the case the *WM* Handshake does not succeed, the user application, based on its configuration, may decide whether to start or not a traditional transaction with the remote application. In fact one of the advantages of this approach is that it does not alter the current Web model.

In our lab, we have built a basic prototype implementing the procedures above in the context of OpenID and applied it to the case of browsing a simple web application.

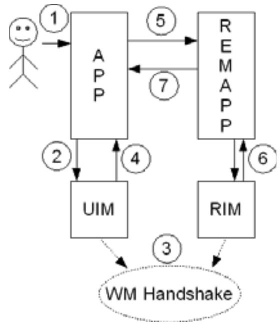


Figure 4. The general architecture.

4. The Online Auction Websites case study

In this section we will show how our architecture can be easily instantiated to a real-life application.

4.1. The case

Alice is an Ebaia power seller with a positive feedback rate of 99%. Thanks to her excellent reputation, Alice reaches big sales volumes. During the Web surfing, Alice finds a new online auction system, called Xbid that offers more convenient commissions on sales. Alice, interested by the offer decides to test the new system but then she finds a serious obstacle: there are no ways to migrate her excellent reputation data (that builds up in a long time span) from the current system to the new one. Discouraged, she decides not to try Xbid.

The adoption of our model, thanks to the relocation of the reputation data on an Identity Provider, allows the user to access to more online auction systems, even at the same time, increasing the seller presence on the market. Also, due to the centralized reputation data, users can compare sellers on different auction platforms allowing a deeper level of filtering. Last but not least, due to the buyers' certified identity, the seller is able to exclude malicious users that can alter the auctions.

4.2. The implementation

The user application is the Web browser (the buyer's one, in this case) and the remote application is the auction system Web server that will request to the seller *RWM* the autorization to proceed with the transaction. The seller *RWM* will be identified by the *UWM* due to a metatag link present in the product page as usually done with *OpenID* delegation. The transaction steps are then so instantiated:

1. the user selects the 'buy now' option;
2. the browser contacts the user *UWM*, through a plugin, to obtain an authorization for the transaction;
3. the *WM* Handshake occurs;
4. if the handshake succeeds, the *UWM* sends the shared *RWM OpenID* authorization token to the browser;
5. the browser sends the 'buy' request together with the authorization token to the auction web system;
6. the auction system uses the token to query its *RWM* to receive the authorization for the incoming request;
7. the auction system shows the payment procedure to the user.

5. Conclusions

In this paper we have presented an architecture for improving some aspects related to Web navigation. The work is still in progress and, due to the complexity of the different addressed issues, many aspects are still to be investigated: some scenarios have been outlined and the architectural model has been presented and tested in one of them. As future work, we plan to test the architectural model in many other scenarios and contexts.

References

[1] A. Ankolekar and D. Vrandečić. Kalpana - enabling client-side web personalization. In *HT '08: Proceedings of the nineteenth ACM conference on Hypertext and hypermedia*, pages 21–26, New York, NY, USA, 2008. ACM.

[2] J. Hoyt, J. Daugherty, and D. Recordon. Openid simple registration extension 1.0. June 2006. http://openid.net/specs/openid-simple-registration-extension-1_0.txt.

[3] A. J, R. Ismail, and C. Boyd. A survey of trust and reputation systems for online service provision. *Decis. Support Syst.*, 43(2):618–644, 2007.

[4] J. Miller. Yadis 1.0. March 2006. http://yadis.org/papers/yadis-v1.0.pdf.

[5] Y. Oiwa, H. Takagi, H. Watanabe, and H. Suzuki. Pake-based mutual http authentication for preventing phishing attacks. In *18th International World Wide Web Conference (WWW2009)*, April 2009.

[6] D. Recordon and D. Reed. Openid 2.0: a platform for user-centric identity management. In *DIM '06: Proceedings of the second ACM workshop on Digital identity management*, pages 11–16, New York, NY, USA, 2006. ACM Press.

[7] D. Reed and D. McAlpin. Extensible resource identifier syntax 2.0 (oasis xri committee specification). November 2005. http://www.oasis-open.org/committees/download.php/15377.

[8] A. C. Squicciarini, A. Trombetta, E. Bertino, and S. Braghin. Identity-based long running negotiations. In *DIM '08: Proceedings of the 4th ACM workshop on Digital identity management*, pages 97–106, New York, NY, USA, 2008. ACM.

The *ENVISION* Project: Towards a Visual Tool to Support Schema Evolution in Distributed Databases

Giuseppe Polese and Mario Vacca

Dipartimento di Matematica e Informatica, Università di Salerno
Via Ponte don Melillo, 84084 Fisciano (SA), Italy
{gpolese, mvacca}@unisa.it

Abstract

Changes to the schema of databases naturally and frequently occur during the life cycle of information systems; supporting their management, in the context of distributed databases, requires tools to perform changes easily and to propagate them efficiently to the database instances. In this paper we illustrate ENVISION*, a project aiming to develop a Visual Tool for Schema Evolution in Distributed Databases to support the database administrator during the schema evolution process. The first stage of this project concerned the design of an instance update language, allowing to perform schema changes in a parallel way [14]; in this paper we deal with further steps toward the complete realization of the project: the choice of a declarative schema update language and the realization of the mechanism for the automatic generation of instance update routines. The architecture of the system, which has been implementing, is also designed.*

1. Introduction

Updating a schema is a very important activity which naturally and frequently occurs during the life cycle of information systems, due to different causes, like, for example, the evolution of the external world, the change of user requirements, the presence of errors in the system. Two of the problems arising when a schema evolves are the semantic of changes (how to express the changes to the schema) and the change propagation (how to propagate the schema changes to the instances) [18]. These two tasks are performed using schema evolution languages and tools. Developing a tool for schema evolution in distributed databases is an important and challenging task for the following reasons: first, the shortage of tools for schema evolution is a well known problem [2, 6]; second, the rare existing tools are limited[1]; changes in distributed database schemas can provoke significant effects because updating instances can involve the processing of an enormous mass of data among distributed nodes, making the process of propagating changes to the instances a very expensive one. As a consequence, database administrators (DBAs) have to cope both with the difficulty of performing schema changes and the efficiency of the change propagation process.

In order to develop a tool, it is necessary to design a schema evolution language, which, according to Lagorce et al., is composed of two languages, the instance update language and the schema update language, and a mechanism allowing to translate schema update statements in instance update ones [11].

The ENVISION (EfficieNt VIsual Schema evolutION for distributed databases) Project aims to develop a Visual Tool to support the DBA during the schema evolution process. The first stage of this project[2] [14] concerned the design of an instance update language, based on the MapReduce Google programming paradigm [5, 7], allowing to perform instance updates in a parallel way [14]. At this stage, the project still suffered from the drawbacks of procedural features of the language.

In this paper we illustrate the second stage of the project, aiming to overcome these problems: we propose to adopt a logical schema update language, both suitable for describing schema changes and straightforward translatable into the instance update one. The result is, hence, the possibility to perform changes to the schema in a declarative way and to let the system generate the MapReduce instance update routines, combining simplicity of use and efficiency.

The paper is organized as follows: after a short introduction to schema evolution and related problems in dis-

[1] For example, the ESRI package ArcGIS (`http://www.esri.com/software/arcgis`) includes tools for geodatabase schema changes, but it supports only a small set of schema changes.

[2] It was developed in collaboration with the Dip. di Costruzioni e Metodi Matematici in Architettura of the Federico II University of Naples.

tributed databases (sections 2 and 3), in section 5 a declarative schema update language is proposed, together with the algorithm for the automatic generation of the instance update routines (introducted in section 4). Section 6 gives an account of the proposed architecture and the conclusions end the paper.

2. Schema evolution: short state of the art and issues

Schema evolution takes place when a schema S evolves towards a schema T (S and T are called schema versions). Two important issues of schema evolution are the management of changes to the schema (a.k.a. semantics of schema changes) and the propagation of changes to the data (a.k.a. change propagation) [18]. The first one refers to the way the changes are performed and their effects on the schema itself, while the second deals with the effects of schema changes on the data instances. These two tasks are realized by schema evolution languages which are, in turn, composed of two languages, the schema update language and the instance update language, and of a translation mechanism allowing to convert schema update statements into instance update ones [11].

Figure 1 describes the schema evolution issues and the role of the schema evolution language: S and T are the two schema versions, I and J are the database instances, map_{ST} denotes a set of statements in the schema update language and instance update routines are the statements by which the database instances are updated accordingly. The big arrow indicates the translation mechanism between the two languages.

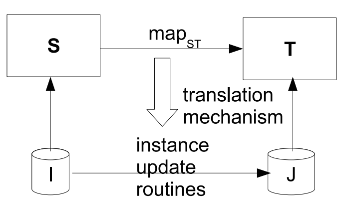


Figure 1. Schema evolution language

According to Lerner [13], there are two classes of schema update languages, differing on the concept of change: the *command approaches* which focus "on the editing *process*" ([13], p. 86) and the ones which focus "on the editing *result*" ([13], p. 86). The approaches belonging to the first class[3] define elementary change operations (like deleting an attribute), by specifying their effects both on the schema and on the data. Changes to the schema can be *simple* (like adding an attribute) or *compound* [13], like merging two relations, which are very important in practical contexts. Two basic features of this kind of changes are their procedural nature and their dependence from the data model. The second kind of approaches are based on the idea that an evolution is a correspondence of schemata (mapping). The first approach of this kind was due to Bertino [3]; the idea was also used in [12] and, later, by Lerner [13] from other points of view. The use of schema mapping to represent schema changes has been increasing more and more, also for the birth of Generic Model Management based approaches, which use schema mappings along with operators to perform schema evolution [2]. Moreover, in a recent research project [6], the Schema Modification Operators have been proposed, whose semantics is expressed by schema mappings. An advantage of the mapping-based approaches is their declarativeness, which makes schema changes easier to realize (e.g. using visual editors).

When a change is applied to a schema, it has to be propagate to the data, either by the DBA [9] or automatically [1, 13]. There are different methods to realize the propagation of schema changes to the instances (see, for example, [13]); in this paper we are interested in the conversion method, where a schema change invokes the update of all the objects affected by the change itself. Two notable examples of instance update languages are those used in the O_2 system [20] and in the TESS system [13].

3. Features and problems of distributed database schema evolution

Distributed databases are applied to a wide variety of domains: from classical administrative databases [18], to e-learning repositories [10] or geographic databases (see, for example, [14]).

There are many kinds of distributed architectures (see [17] for a detailed account), all of them sharing the feature that data are fragmented across (geographically) distributed nodes. In this paper we are interested in all the cases where a central node manages a schema and the data of the database are spread across the local nodes, using fragmentation criteria [17]. The hypotheses of the presence of a central node is not a limitation, as this situation is true for a large number of architectures (for an example, see the POOL architecture [10]).

The interest of schema evolution research in distributed databases has growing in the latest years, as the inclusion of this topic in the most updated survey on schema evolution shows [18]. Within the context of distributed databases, the schema evolution issues of section 2 become more challenging: first, the change propagation process, involving potentially enormous mass of data distributed across nodes, is very expensive, and it calls for efficient processing; second, the translation mechanism is more difficult because the up-

[3] See [1, 19, 20] for taxonomies of schema change operations.

dating routines are more complex. Therefore, the need for a supporting tool, which allows the DBA to formulate the schema changes easily and to propagate them to the data automatically and efficiently, becomes more and more urgent.

4. A MapReduce-based instance update language for distributed databases

MapReduce is a programming model [7] developed by Google to support parallel computations over vast amounts of data on large clusters of machines. The MapReduce framework is based on the two user defined functions *map* and *reduce* and its programming model is composed of many small computations using these two functions. In general, the *MapReduce execution process* (see [7] for details) considers special just one of the copies of the user program calling map-reduce functions (called master), while the rest are workers (there are M mappers and R reducers) the master assigns work to.

The MapReduce model has been extended for processing heterogeneous datasets [5] and it is based on three user defined functions (map, reduce and merge) with the following semantics (see [5] for details): a call to a map function processes a key/value pair $(k1, v1)$ returning a list of intermediate key/value pairs $[(k2, v2)]$; a call to a reduce function aggregates the list of values $[v2]$ with key $k2$ returning a list of values $[v3]$ always with the same key; a call to a merge function, using the keys $k2$ and $k3$, combines them into a list of key/value $[(k4, v5)]$. Notice that a merge is executed on the two intermediate outputs $((k2, [v3])$ and $(k3, [v4]))$ produced by two map-reduce executions.

In [14], the Map-Reduce-Merge model has been exploited as an instance update language for geodatabases. The proposed execution process, inherited from [5, 7], is the following:

- *Map task*

When a map is encountered, the master assigns the map tasks to the M workers (mappers). A map task consists in reading data from the input locations, passing them to the user map function and, then, storing them, sorted by the output key, at some locations on some nodes.

- *Reduce task*

The master passes the locations where the mappers have stored the intermediate data to the R reduce workers (reducers) which are assigned to some nodes. The reducers, using an iterator, for each unique intermediate key, pass both the key itself and the corresponding list of values to the user's reduce function. The result of the user reduce function is stored on some nodes.

- *Merge task*

When the user program contains a merge call, the master launches the merge workers (mergers) on a cluster of nodes, which take data, to be passed to the user merge function, from two sources (the locations where reducers stored them) using both a partition selector and an iterator.

Example 1 *Consider the schema S storing information about cities*

$$S = \{Cities(city, prov, pop), Provinces(prov, reg)\}$$

and the schema T obtained from S joining its relations on the attribute prov:

$$T = \{NewCities(city, prov, pop, reg)\}$$

The instance update related to this change can be realized by the following sequence of map, reduce and merge routines:

```
use input Cities;
map(const Key& key,
    const Value&, value){
    prov = key;
    city = value.city;
    pop = value.pop;
    Emit(key,value);
}
/* This map reads the Cities tuples from
the input locations and stores them, sor-
ted by the output key prov, at some loca-
tions on some nodes.*/

reduce(const Key& key,
       const Value& value){
 Emit(key, value);
}
/*This reduce function, for each unique
intermediate key prov, builds the corre-
sponding list of values.*/
use input Provinces;
map(const Key& key,
    const Value&, value){
    prov = key;
    reg = value.reg;
    Emit(key,value);
}
/*Analogous to the previous map.*/
reduce(const Key& key,
       const Value& value){
 Emit(key, value);
}
/*Analogous to the previous reduce.*/
merge(const LeftKey& leftKey,
      const LeftValue& leftValue,
      const RightKey& rightKey,
      const RightValue& rightValue) {
 if (leftKey == rightKey){
 Emit(leftKey,rightKey);}
}
```

```
/*The merge joins the result of the two
previous reduce functions on prov.*/
use output NewCities;
divide NewCities;
/*The table NewCities is fragmented.*/
```

5. The schema evolution language

The features of distributed database schema evolution of section 3 lead to the following requirements for the schema evolution language[4]:

- the language to express schema changes has to be declarative, possibly visual;
- the mapping between schema versions has to have a formal (logical) characterization;
- instance update (MapReduce-based) routines have to be generated automatically;
- it has to be always possible to choose the level to operate with: visual (schema mapping) or instance.

The independence of use of the instance update language from the visual schema update (the DBA has to be free to choose any of them) is particularly important, as very complex schema changes could be required which are not supported, or not efficiently enough, by the tool.

5.1. The schema mapping language

An important problem to cope with when designing the schema evolution language is the choice of the formal language for the mappings between schema versions. Mappings link two schemas S and T and are represented by "set of formulas of some logical formalism over (S, T)" (Fagin et al. [8], p. 999) describing the relation between the instances of the two schemas themselves (see Figure 1). There are many logical schema mapping languages (see, for example, [16] for a list), each of them suitable for some purposes. Among them, the *second order tuple-generating dependency (SO tgd)* language [8] has many desirable properties: it allows to express many schema changes (note that SO tgd class includes that of GLAV mappings, which are sufficient to link schemas for practical goals [15]); it has been proved to be closed under composition [8]; its statements can be easily decomposed (see [8]); it allows the use of functions. Moreover, we will show that its statements can be also easily translated in MapReduce based instance update routines.

A second order tuple-generating dependency (SO tgd) (see [8] p. 1014 for details) is a formula of the form:

$$\exists \mathbf{f}(\forall \mathbf{x}_1(\phi_1 \rightarrow \psi_1) \wedge \ldots \wedge \forall \mathbf{x}_n(\phi_n \rightarrow \psi_n))$$

where $\mathbf{f}$ is a set of functions, ϕ_i (resp. ψ_i) $(i = 1, \ldots, n)$ is a conjunction of atomic formulas of the form $S_j(y_1, ..., y_k)$ (resp. $T_j(y_1, ..., y_k)$), with S_j (resp. T_j) k-ary relations of S (resp. T) and $y_1, \ldots, y_k$ variables in x_i (resp. terms on x_i and f).

The language we propose to use is based on SO tgds, but, since we use it in practical applications, we need instantiated SO tgds formulas (we call them *ISO tgds*): first, the set of function $\mathbf{f}$ has to be instantiated (the DBA has to write them, if necessary); second, in order to perform the join (the right side ϕ of a SO tgd is a conjunction), the DBA has to specify the merge attributes (this is made using *equality constraints* stating what attributes have to be considered equals; if no constraint is specified, the attributes with equal name and type are considered equal, and if none of such attributes exists, the join is interpreted as cross join).

Definition 1 *Let S and T be two schemas. An ISO tgd mapping is a triple (Σ, E, F), where Σ is a set of SO tgds, E is a set of equality constraints on S, and F is a set of assignments of the kind y = f(x) (f is a function, x is a list of attributes of relations in S and y is an attribute of some relation in T).*

A simple example of function is the one assigning default values when a column is added to a table.

Example 2 *Consider the schema evolution in the example 1. The ISO tgd mapping describing the passage between the two schema versions S and T is:*
$\Sigma = \{\forall city, prov, pop, reg($
$Cities(city, prov, pop) \wedge Provinces(prov, reg) \rightarrow$
$\rightarrow NewCities(city, prov, pop, reg))\}$
$E = \{Cities.prov = Provinces.prov\}$ *and* $F = \emptyset$.

5.2. The translation mechanism

Even if the Map-Reduce-Merge language is procedural, it is just its basic feature (i.e. being based on only three functions) to suggest the possibility to generate instance update routines automatically. The idea under the automatic generation is to use "basic" routines (we call them *propagator chunks*) which, properly combined, generate the desired instance update ones.

Definition 2 *(Propagator chunks) Let S be a relation and let $[y_1, \ldots, y_n]$, k be, respectively, a list of attribute names and an attribute name (a key); let $[f_1(\mathbf{x}_1), \ldots, f_n(\mathbf{x}_n)]$ be a list of function names f_i, each with its argument name list x_i $(i = 1, \ldots, n)$:*
- ***map-chunk***$(R, k, [y_1, \ldots, y_n])$ *is the routine:*

```
use input R;
map(const Key& key,
    const Value&, value){
```

[4] These desiderata are a remake of Curino et al. D1.1, D1.4, D3.4, D3.7 ones [6].

```
    k = key;
    y1 = value.y1;...yn = value.yn;
    Emit(key,value);
}
```

The map-chunk reads the data from the input locations of the table R and stores the values of the attributes $k, y_1, \ldots, y_n$, sorted by the output key k, at some locations on some nodes.
- ***reduce-chunk***$([y_1, \ldots, y_n], [f_1, \ldots, f_n])$ *is the routine:*

```
reduce(const Key& key,
       const Value& value){
 y1 = f1(x1);...yn = fn(xn);
 Emit(key, (y1,...,yn));
}
```

Moreover, if $[y_1, \ldots, y_n]$ is empty, the reduce-chunk ends with Emit(key,value) instead of Emit(key, $(y_1, \ldots, y_n)$) and if f_i = nil (the no-operation function), there will be no assignment $y_i = f_i(x_i)$.
The reduce-chunk, using an iterator, for each unique intermediate key k, pass the list of values to the user reduce functions $f_1, \ldots, f_n$; the result of the user reduce functions is stored on some nodes.
- ***merge-chunk***(E) *is the routine:*

```
merge(const LeftKey& leftKey,
      const LeftValue& leftValue,
      const RightKey& rightKey,
      const RightValue& rightValue)
 if (E) {
 Emit(leftvalue,rightvalue);}
}
```

The merge-chunk takes data from two sources (the locations where reducers stored them) and merges them using the set E of equality constraints.
- ***divide-chunk****(R) is the routine:*

```
use output R;
divide R;
```

The divide-chunk(R) fragments table R across the nodes.

The following *IURG* algorithm is an instance update routine generator, using the propagator chunks.

```
Algorithm IURG(Σ,Ω);
INPUT   an ISO tgd mapping (Σ, E, F);
OUTPUT  the set Ω of Map–Reduce–Merge
        routines ρ;
begin
 Σ' := ∅;
 for each σ ≡ φ → ⋀_{i=1}^{n} T_i ∈ Σ begin
     add σ_i ≡ φ → T_i(z) (i = 1,...,n) to Σ';
 end;
 K_1 := ∅; K_2 := ∅;
 for each σ in Σ' do begin
 {σ has the form φ → T(z)}
   ρ := Λ; {ρ is set to the empty string}
   for each S(y) in φ do begin
     update the key sets K_1 and K_2 using E;
     add map–chunk(S,K_2,y) to ρ;
     add reduce–chunk([],[]) to ρ;
     if S is not the first relation in φ then
        add merge–chunk(E_{K_1,K_2}) to ρ;
           {E_{K_1,K_2} is set of constraints in E
            restricted to K_1 and K_2}
     if S is the last relation in φ then
        add reduce–chunk(z,F) to ρ;
   end;
 add divide–chunk(T) to ρ;
 Ω = Ω ∪ {ρ};
 end;
end {IURG}.
```

It is easy to see that the computational complexity of the algorithm is $\mathcal{O}(|\Sigma'| \cdot max|\phi|)$, where $max|\phi|$ denotes the maximum number of relation symbols in a right side formula ϕ.

Example 3 *The IURG algorithm, applied to the ISO tgd in the example 2, produces the instance update routine generated by the following list of propagator chunks:*
map-chunk(Cities,prov,[city,pop])
reduce-chunk([],[])
map-chunk(Provinces,prov,[reg])
reduce-chunk([],[])
merge-chunk(E)
reduce-chunk([city,prov,pop,reg],[nil])
divide-chunk(NewCities)

6. The architecture

The architecture of the system, still under development, showed in Figure 2, is constituted by the following modules:

- *Visual Schema Manager* (**VSM**)

  This module is constituted of the **visual interface** (**VI**) and of the **VisualToMapping** translator, which generates the SO tgds associated to visual changes. The visual interface we are realizing is inspired from the famous Clio project [15] and it allows to create mappings between schema versions using visual operators like select, link, move, delete, add and modify. It also allows to write functions on attributes to be associated to other attributes, and to specify equality constraints.

- *Instance Update Routine Generator* (**IURG**)

  This module, based on the IURG algorithm presented in section 5.2, takes a SO tgd as input and returns the Map-Reduce-Merge instance update routines.

- *Network Manager* (**NM**)

  This module coordinates the execution process described in section 4. It also provides an interface to write Map-Reduce-Merge routines.

The system uses the Java platform and hadoop.

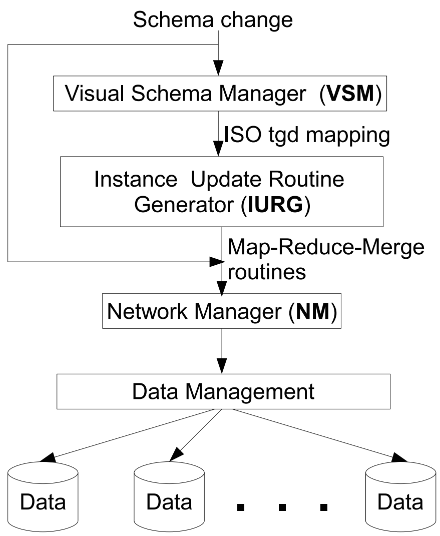


Figure 2. The *ENVISION* system architecture

7. Conclusions and future work

A schema update language, together with an algorithm to translate its statements into Map-Reduce-Merge instance update ones, has been presented. This language allows to design a visual interface and, hence, to lay the foundations for building a complete tool to support schema evolution in distributed databases, whose architecture has also been presented. The next step we have planned is to enrich our model with a simulation function (extending the **NM** module functions) to check the change effects before performing them: on the one hand, this provides the DBA with a further tool to manage changes, and, on the other hand, such a function is a very important tool for us in order to study the efficiency of the system, that is to fulfill our goal of making the schema evolution process as much efficient as possible in distributed databases.

References

[1] J. Banerjee, W. Kim, H.-J. Kim, and H. F. Korth. Semantics and implementation of schema evolution in object-oriented databases. In U. Dayal and I. L. Traiger, editors, *SIGMOD Conference*, pages 311–322. ACM Press, 1987.

[2] P. A. Bernstein and S. Melnik. Model management 2.0: manipulating richer mappings. In Chan et al. [4], pages 1–12.

[3] E. Bertino. A view mechanism for object-oriented databases. In A. Pirotte, C. Delobel, and G. Gottlob, editors, *EDBT*, volume 580 of *Lecture Notes in Computer Science*, pages 136–151. Springer, 1992.

[4] C. Y. Chan, B. C. Ooi, and A. Zhou, editors. *Proceedings of the ACM SIGMOD International Conference on Management of Data*. ACM, 2007.

[5] H. chih Yang, A. Dasdan, R.-L. Hsiao, and D. S. P. Jr. Map-reduce-merge: simplified relational data processing on large clusters. In Chan et al. [4], pages 1029–1040.

[6] C. Curino, H. J. Moon, and C. Zaniolo. Graceful database schema evolution: the prism workbench. *PVLDB*, 1(1):761–772, 2008.

[7] J. Dean and S. Ghemawat. Mapreduce: Simplified data processing on large clusters. In *OSDI*, pages 137–150, 2004.

[8] R. Fagin, P. G. Kolaitis, L. Popa, and W. C. Tan. Composing schema mappings: Second-order dependencies to the rescue. *ACM Trans. Database Syst.*, 30(4):994–1055, 2005.

[9] F. Ferrandina, T. Meyer, R. Zicari, G. Ferran, and J. Madec. Schema and database evolution in the o2 object database system. In U. Dayal, P. M. D. Gray, and S. Nishio, editors, *VLDB*, pages 170–181. Morgan Kaufmann, 1995.

[10] M. Hatala and G. Richards. Global vs. community metadata standards: Empowering users for knowledge exchange. In I. Horrocks and J. A. Hendler, editors, *International Semantic Web Conference*, volume 2342 of *Lecture Notes in Computer Science*, pages 292–306. Springer, 2002.

[11] J.-B. Lagorce, A. Stockus, and E. Waller. Object-oriented database evolution. In F. N. Afrati and P. G. Kolaitis, editors, *ICDT*, volume 1186 of *Lecture Notes in Computer Science*, pages 379–393. Springer, 1997.

[12] L. V. S. Lakshmanan, F. Sadri, and I. N. Subramanian. On the logical foundations of schema integration and evolution in heterogeneous database systems. In *DOOD*, pages 81–100, 1993.

[13] B. S. Lerner. A model for compound type changes encountered in schema evolution. *ACM Trans. Database Syst.*, 25(1):83–127, 2000.

[14] F. D. Martino, G. Polese, S. Sessa, and M. Vacca. A mapreduce framework for change propagation in geographic databases. In *ICEIS*, 2009.

[15] R. J. Miller, M. A. Hernández, L. M. Haas, L. Yan, C. T. H. Ho, R. Fagin, and L. Popa. The clio project: managing heterogeneity. *SIGMOD Rec.*, 30(1):78–83, 2001.

[16] A. Nash, P. A. Bernstein, and S. Melnik. Composition of mappings given by embedded dependencies. *ACM Trans. Database Syst.*, 32(1):4, 2007.

[17] M. T. Özsu. Distributed database systems. In *Encyclopedia of information systems*, pages 673–682, 2003.

[18] S. Ram and G. Shankaranarayanan. Research issues in database schema evolution: the road not taken. *Univ. of Arizona, Working Paper #2003-15*, 2003.

[19] J. F. Roddick, N. G. Craske, and T. J. Richards. A taxonomy for schema versioning based on the relational and entity relationship models. In *ER*, pages 137–148, 1993.

[20] R. Zicari. A framework for schema updates in an object-oriented database system. In *ICDE*, pages 2–13. IEEE Computer Society, 1991.

Towards Synchronization of a Distributed Orchestra

Angela Guercio
Department of Computer Science,
Kent State University Stark
North Canton, OH 44720, USA
e-mail: aguercio@kent.edu

Timothy Arndt
Dept. of Computer and Information Science
Cleveland State University
Cleveland, OH 44115, USA
e-mail: arndt@cis.csuohio.edu

Abstract
In an Internet-based multimedia application that plays an orchestra of remote source sounds, the synchronization of audio media streams is essential for optimal performance of the piece. The application that enables this virtual synchronized orchestra benefits from the use of a language containing constructs that help express the specifications and requirements of such a reactive system. We provide a model for the performance of a distributed orchestra. The architecture of the conducting system takes advantage of the synchronization abilities of TANDEM, a multimedia language for reactive multimedia systems that has been extended with constructs to describe the conductor's gestures and the syntax and semantics of those constructs. The PCM live streams and at least one MIDI stream per section are multiplexed at each remote source and time stamped before transmission. At the receiver the TANDEM environment performs synchronization with the trigger and the active repository.

Index Terms – **Computer Languages, Multimedia Systems, Real Time Systems, Synchronization, Reactive Systems.**

1. INTRODUCTION

Music has been widely used to entertain, relax (in doctors offices, elevators and commercial centers), and nourish the artistic spirit. The ability to download on a PDA or MP3 player our favorite orchestra piece is a reality. The popularity of tools like YouTube, iPhone and iPods are examples and industry and research have devoted much attention to multimedia tools which increase our ability to interact with media to communicate. All these multimedia tools require a type of synchronization in order to produce the desired outcome. We focus our attention on the problem of an Internet-based multimedia application that plays an orchestra composed of distributed sounds.

The possibilities when distributed remote audio streams are synchronized together are endless. To mention a few:

a) the creation of a Virtual Orchestra with sound tracks coming from remote sources would be an invaluable tool for a musician who wants to experience the execution of a piece with his/her favorite musician;
b) the ability to create extemporaneous virtual sonority that can be added to other media types, to recreate the sound of a specific environment in a museum (e.g., the sounds of the savannah in Africa at dawn);
c) the live performance with musicians playing in different parts of the world.
d) in classrooms, to add to the local students' musical performance, a remote soloist.

The strong synchronization required by distributed musical applications can be beneficial in applications in the domains of distance education, large-scale military training, homeland security, business or social meetings.

The result of the orchestra performance must be a realistic reproduction of the composer's beat, tempo, and expression symbols performed in a synchronized way that avoids possible stuttering effects or unsynchronized performance. The satisfaction of all these requirements is challenging, and has led to the development of special-purpose languages for multimedia authoring and presentations. In particular for computer techniques applied to music and musicology which deal with audio and/or graphical representation or score of music with performance and sometimes with choreography, several recommendations or standards have been introduced. Examples of such languages include the latest IEEE 1599 standard [15], SMIL [19] and all the existing markup music initiatives [21] such as SMDL, MusiXML, MusicXML, MDL, FlowML, Hy Time, etc.. While some of the above languages only describe musical notation, others can describe a multimedia presentation containing multiple media sources, both natural and synthetic, as well as stored or streamed media. In SMIL and IEEE1599 some mechanisms for specifying layout of the media on the screen is given as well as primitives for synchronizing the various elements of the presentation and a small set of basic events are supported while more complex events require the use of scripting languages such as JavaScript.

While these languages are well suited for the description of music and multimedia presentations on the Web, they are of limited use for creating more general distributed multimedia applications since general-purpose programming is only available through scripting languages that have limited power. To support the construction of more large-scale applications approaches such as the use of special multimedia libraries along with a language as in the case of Java and JMF [10] or extension of middleware such as CORBA [17] are available. Besides lacking essential characteristics for development of advanced distributed multimedia applications that will be noted below, the use of libraries and/or middleware to achieve synchronization and perform other media related services results in a less well-specified approach than can be achieved by directly extending existing general purpose languages with multimedia constructs with precisely specified semantics.

Following this approach, in [5] a language, called TANDEM (Transmitting Asynchronous Non-deterministic and Deterministic Events in Multimedia systems) and its architectural model [6, 7, 8] that supports general-purpose computation has been analyzed and designed. The language constructs can be added to an existing general purpose language such as C, C++ or Java. This approach is similar to the approach taken by the reactive language Esterel [2, 1] which adds reactivity to general purpose languages. We extend the language by introducing the syntax and semantics of new constructs for synchronization of a distributed orchestra of audio media. These constructs express the temporality of the piece as derived from the conductor's gesture. The semantics of these constructs expresses the temporal issues required and enforces the generation of appropriate events. The TANDEM architectural model is able to deal with audio streams so that they can be played in temporal correlation, to guarantee the synchronization after possible transformations (e.g. transpositions, distortions, etc.) and to handle possible data loss during transmission over the channel.

2. THE GESTURES OF THE CONDUCTOR

The conductor of a live orchestra is a simple time-keeper as well as an interpreter and communicator of emotional content of the music being played. Classical studies on the conductor's gesture can be found in manuals such as [16]. While different conductors direct the orchestra according to their personality and expressivity, this should not affect the pure synchronization aspect of the final execution but rather increases the beauty of the performance for the listener.

According to Luck, [13, 14], who has performed an empirical investigation, no effects of the conductor's previous experience or the radius of curvature with which the beat was defined alter the conductor-musician synchronization. Only the experience of the participants in the experiments was significant and affected their synchronization ability. On the basis of these results, we assume for simplicity a set of conducting gestures such as [18]. The gestures are independent of the experience of the conductor. Each gesture of a conductor is represented by a vector. A gesture has a speed measured by a quantum.

We assume that the conductor performs with two hands, one hand maintains the beat (we assume the right), the other controls volume, attack etc. In particular, for the left hand,

a) the *vertical* gesture of the conducting hand down or up controls the volume, which increases when the direction goes up and decreases when it goes down. This gesture is interpreted as a *crescendo* or *diminuendo* during the piece execution.
b) the *horizontal-toward* gesture (horizontal towards a section of instruments moving the hand in a downward movement) of the conducting hand lets the conductor start the section's audio stream. This gesture is interpreted as an *attack*. The consequence of starting a section by pointing at it starts the buffering of the section's audio stream while the playback of the section's stream will start according to the synchronization of the beat.
c) the *circular* gesture of the conducting hand lets the conductor bring a section or the whole orchestra to a stop. This gesture is interpreted as an *interrupt*.

When a gesture starts, the speed, direction and amplitude are identified.

The speed of the gesture: The speed is assumed to be maintained constant for the duration of a gesture. The speed of the gesture is used to help the synchronization of the multimedia streams with the beat.

The direction of the gesture: The direction of the gesture is used to identify the gesture and when the change of gesture occurs. The direction may also identify the change of volume, as in the Vertical gesture, or an attack, as in the Horizontal-Toward gesture.

The amplitude of the gesture: The amplitude of the gesture is important in the vertical gesture since the larger the amplitude of the gesture, the higher the volume of the orchestra. The change in volume is never abrupt and is modeled by a progressive variation.

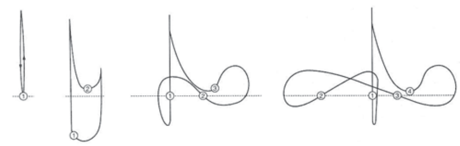


Fig. 2.1 Orchestra Conductor Movements

In the case of a virtual orchestra with independent remote and local sources, multimedia data integration and presentation must take into consideration the synchronization of the media streams that are streamed in real-time and, possibly compensate for jitter or any other possible alterations caused by the orchestra participants.

3. THE PERFORMANCE ENVIRONMENT

In this section we will give an overview of the performance environment. The principal actors are the orchestra sections, each composed of a group of performers. Each section may contain either live and/or recorded musicians, local or remote. The other principal actor is the conductor who may be either a live or virtual conductor. Anything that is local is in the same location as the conductor. If TANDEM is extended with commands to describe the conductor's gestures, a *virtual conductor* can be used to create an animated avatar representing the conductor and reproducing the correct gestures. The system will respond to the gestures and will produce reactions to various situations using triggers and the active repository.

In order for the conductor and live musicians to interact, the actors must each be able to see each other. The musicians see the conductor by viewing a video stream (either of the live or animated virtual conductor). The conductor needs to have a view of the entire orchestra to direct his gestures at particular sections. This is done through a *virtual stag*e (see fig. 3.1) - a screen with multiple windows each containing one or more sections. The position of these windows/sections is defined using spatial relations as in SMIL. When the performance begins, the conductor sees a screen containing the virtual stage. He can then direct his gesture to the relevant sections. Each window on the virtual stage can be filled either with a live or recorded video of the section, or with a static image or animated representation of the section.

The live conductor's gestures are captured using gesture recognition techniques, possibly incorporating sensors [20] or computer vision technology [3]. The gestures are then transformed into conductor actions (see the following section) using motion tracking and classification algorithms. The conductor's actions can be used to drive an animation of the conductor for remote live performers if a video stream of the conductor is not available.

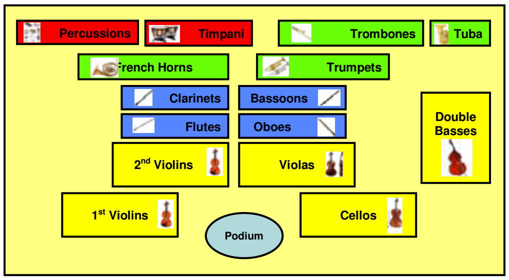


Fig. 3.1 The Virtual Stage

Live local performers respond immediately to the conductor's gestures, while remote performers response is somewhat delayed due to differing amounts of network latency. This difficulty is overcome by the Active Repository [9] which acts as a buffer for remote and recorded performance data. Immediate response to the conductor's gestures is achieved via synchronization constructs of TANDEM on the data in the Active Repository (see fig.3.2).

Just as live performance on MIDI devices can be captured for later playback, so can the conductor's gestures be captured (in the Active Repository) and later "played back", that is, used to drive a virtual conductor animation or avatar to control an orchestra. Also analogously to MIDI, it is possible to program the conductor's performance (without going through the actual conducting gestures – analogous to composing MIDI scores without actual performance) and use the program to drive the virtual conductor. Of course it is usually necessary for the conductor to respond to the performers during the performance. This can be supported by the system by defining a number of triggers based on performance conditions which will be triggered by meeting the conditions and cause particular gestures to be generated/modified.

The structural model of the system is depicted in fig. 3.3.

4. *Language Constructs*

In this section we will describe the language constructs that support synchronization for a synchronized orchestra. The exact syntax of the constructs will depend on the host language the multimedia constructs are embedded in. In the examples that are given, the host language is C. This results in the constructs having a "C-like" syntax. It is expected that the processing of the synchronization constructs will be handled by a preprocessor before passing the results to a compiler for the given host language. A run-time environment then supports run-time synchronization. The implementation will be similar to Esterel [1].

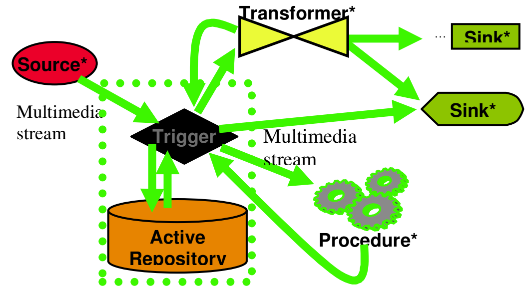


Fig. 3.2 The TANDEM Model

A. Construct Definition

The syntax of the constructs must express the synchronization between the gestures of the conductor and the multimedia streams that represent the instruments of the orchestra. The conductor uses a virtual stage interface where the sections have been spatially arranged on the screen before the beginning of the performance. On the virtual stage a section represents a number of musicians grouped by instrument type, the i.e. 1st violins, the 2nd violins, the flutes, and so on. A section consists of a set of one or more media streams or it could be under local control. Multiple musicians playing together remotely will be captured by a single camera and a single media stream will be transmitted over the network. Multiple musicians remotely located in multiple geographic locations that are part of the same section produce a number of streams equal to the number of remote locations. Multiple sections of the virtual orchestra may also be multiplexed together in a single stream if they are at the same remote location. Multiple musicians of the same section that play locally are not associated with any stream.

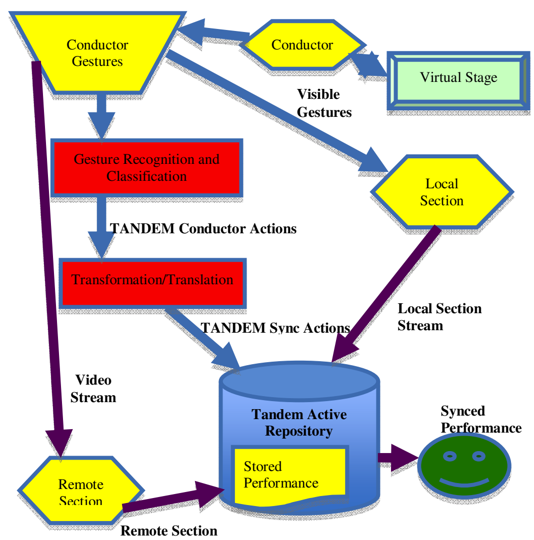


Fig. 3.3 The Structural Model of the Conductor System

Conductor gestures may be directed at the orchestra as a whole or at individual sections. Gestures directed at individual sections may be classified as either *immediate local; immediate remote;* or *delayed remote*. The gesture is immediate local if it directed at a local section. The gesture is immediate remote if it is directed at a remote section that produces one multiplexed stream. In this case, the remote players of the section will require a certain, small amount of time (depending on the roundtrip network latency) to respond to the gesture, but the Active Repository can mask this latency by modifying the buffered stream (increasing playback rate, decreasing volume, etc.). If the gesture is directed at a part of the remote section (this would occur if a remote section contains more than one instrument type) since the remote section produces one multiplexed stream, the latency in responding to the gesture by part of the remote section cannot be masked by the Active Repository, since this would involve modifying (e.g. speeding up) multiple sections of the multiplexed stream, not just the single one to which the gesture is addressed. We assume that the gestures intended for the virtual orchestra as a whole are either immediate local or immediate remote (there exists one or more remote sections and each remote section generates a multiplexed stream).

Given the previously defined virtual stage, we define the virtual orchestra as a group of pairs section/region.

```
group my_orchestra = (section1, region1,
            section2, region2 …)
```

Each instrumental section then must be defined as associated with one or more streams or as a local section. For example:

```
section wind =(windstream1, windstream2)
section chorus =(local)
```

The streams are defined in TANDEM in terms of their various attributes. The actions of the conductor are connected with the gestures recognized by the gesture analyzer that the conductor performs to guide the sections of the orchestra. The enumerated list of available actions is:

```
enum actions {beat, attack, interrupt,
      cutoff_section, cutoff, crescendo,
      diminuendo}
```

The time signature is indicated by the “beat”. The beat is given by the gesture of the right hand of the conductor. The beat is identified by the change of direction of the end of the baton. The gesture analyzer produces the command

```
beat(time, position)
```

The speed of the baton can be derived from the times and positions of a sequence of two beats. We assume that the conductor, as well as the musicians, are aware of the time signature of the piece being performed. The rigid value of the metronome can be slightly stretched by the personality of the conductor which can be detected through the change of speed between beats. At each beat the synchronization of the media streams is enforced.

The *attack* gesture indicates that a section or the orchestra as a whole should start to play. The gesture is a horizontal-toward gesture (pointing) with the left hand directed at the section or orchestra. Both the time of the attack and section indicated or orchestra as a whole are retrieved and passed as parameters to the command by the gesture analyzer. The command is described as:

```
attack(time, section)
attack(time, orchestra)
```

The time of the attack is synchronized with the time of the beat relative to the stream indicated by the section parameter. If the attack is directed to the whole orchestra, all the streams will be synchronized as a group.

We assume that when two sections are addressed to start at the same time two sequential movements indicating attack are detected in very close time sequence. Such time difference is smaller than an ε (the ε must be smaller than a beat time) the two attacks are interpreted as one, the two sections are processed as being in one group and the multiplexed streams relative to the two involved sections are synchronized with respect to the first common synch point detected among the group participants. In a fine synchronization, the distance between synch points must be imperceptible to the human ear.

The crescendo (resp. diminuendo) gesture which is indicated by an upward (resp. downward) vertical movement of the open left palm, increases (resp. decreases the volume of a section. The command produced by the gesture analyzer contains the time at which the gesture occurs as well as the section to which it applies or the orchestra as a whole and is described as:

```
crescendo(time, section)
crescendo(time, orchestra)
```

(resp., `diminuendo(time, section)`
`diminuendo(time, orchestra)`)

The semantics of the command enforces the volume alteration accordingly in a synchronized way.

The circular gesture of the left hand is used to interrupt a section or the whole orchestra. The command produced by the gesture analyzer contains the section to which it applies and is described as:

```
interrupt(time, section)
interrupt(time, orchestra)
```

The command causes an abort of the section or whole orchestra.

5. THE SEMANTICS OF THE CONSTRUCTS.

The conductor system is a distributed multimedia reactive system modeled as a communicating concurrent state machine in which multiple triggers are concurrently active at different remote sites (see fig. 3.2). We distinguish two types of states: a *computational* state and a *multimedia* state. A *computational* state is a set (*identifier, attribute, value*), where an entity could be a stream, variable, constant, spatial constraint, temporal constraint, mobile processes caused by migration of code over the Internet, or a channel between two computational units. A *multimedia state* M is a set of multimedia entities such as streams, asynchronous signals (denoted by η), partial conditions, or attributes of media objects such as streams or asynchronous signals. A transition between multimedia states occurs if media entities are transformed. A change of multimedia state also generates changes in computational states. Transformation of a multimedia state involves passage through many computational states with no multimedia state change.

In a real time conductor system, two concepts are very important: *continuity,* which contains the notion of temporality, and *context* which expresses spatio-temporal relationship between objects. Breakage of either of them causes lack of perception and comprehension. In multimedia reactive systems, continuity is guaranteed by the physical presence of the multimedia streams and their temporal relationship to each other which is guaranteed by the presence of multiple clocks and the presence of synch points at regular temporal intervals. The temporal logic of the system and the state behavioral semantics provides the behavioral rules for the language by describing the states and transitions between states during computation.

We use state logical behavior to describe the semantics of the constructs introduced. The constructs *attack* and *interrupt* produce events that generate trigger operations. Let α be the action taken that transforms the multimedia state μ into μ', then a state transformation caused by an action α given the set of entities Ψ is written as

$$\mu \xrightarrow[\alpha]{\Psi} \mu'$$

We define a *streaming code number k,* where $k \geq 1$. The streaming code number encodes the reaction to an asynchronous signal, such as attack or interrupt, performed on the streams samples between two synch points of a stream. When k=1 the action of abortion is strong; for k>1 the action abortion is weak. We will denote the state after applying the actions in a single iterative cycle as μ^I. Under the assumption that the smallest data unit is an audio sample or a video frame, the sync point for an audio/generic media stream corresponds to m ($m \geq 1$) data units. Then the state transition for traversing one sync point is $(\alpha^I)^m$. An asynchronous signal η that initiates a preemptive action, such as an *interrupt*, has to wait 1 synch point to reach the new state μ'. However, if the abortion is strong (this is the most general case in the conductor system) the streaming will involve the whole orchestra and is interrupted at the first synch point (k = 1) of the stream and the control moves out of the beat loop. If the interrupt is weak (useful for more general use in multimedia systems) the streaming is completed after the current clip/audio stream is over ($k \geq 1$).

During abortion the current state is saved. However, the multimedia state is defined as the disjoint union of the frozen state and the new state derived from the alternate thread of activity so that the frozen state can be restored after the next *attack* action. At the first attack of the performance there are no frozen states and $\mu^{susp} \oplus \mu' = \mu'$.

Table 1 describes semantic rules for interrupt, and attack.

interrupt	$\dfrac{\mu \xrightarrow[(\alpha^I)^m,\ k \geq 1]{\psi} (\mu^I)^{m*k}}{\mu \wedge \neg\eta \xrightarrow[\text{interrupt};\ (\alpha^I)^m,\ k \geq 1]{\psi} (\mu^I)^{m*k} \wedge \eta}$
suspend	$\dfrac{\mu \xrightarrow[\alpha]{\psi} \mu'}{\mu \xrightarrow[\text{suspend};\ \alpha]{\psi} \mu \oplus \mu'}$
attack	$\mu^{susp} \oplus \mu' \xrightarrow[\text{attack}]{\psi} \mu^{susp}$

Table 1. Semantics of the constructs in the trigger

The constructs *crescendo* (resp. *diminuendo*) perform transformation actions which are executed in the transformer. The construct increase or decrease the volume.

A *stream s* is a pair of the form (s^A, s^D) where s^D is a sequence of elements containing the data and s^A is the set of attributes associated with the stream s. We use $\sigma(s^D, i)$ to denote the i[th] frame/sample (data_element) in the stream. Accessing a frame/sample f in a stream s, is performed by the access operator that is defined as $\pi_1(\sigma(\pi_2(s), i))$ if $0 < i \leq \|s\|$ otherwise $\perp$ (read *undefined)* where π_1 accesses the attribute elements of the stream, and π_2 accesses the data elements of the stream. Therefore, the *crescendo* construct is expressed as $crescendo(s) = \pi_1(\sigma(s^D, i))$.

6. SYSTEM DESIGN AND SYNCHRONIZATION

The system architecture is depicted in fig. 5.1. Each remote source has several musical instruments and one or more MIDI instruments. The PCM audio of the instruments is mixed onsite. The sampled PCM data are multiplexed with MIDI data and stored in time stamped packets. Each packet (see fig 5.2) contains a sequence of PCM samples, followed by a sequence of MIDI events occurring in the time interval, plus a time stamp. The number of samples collected in each packet and the sampling rate, give the granularity of future synchronization.

Header

Time Stamp Number of MIDI events	PCM Samples	MIDI Events

Fig. 5.2 Multiplexed packet

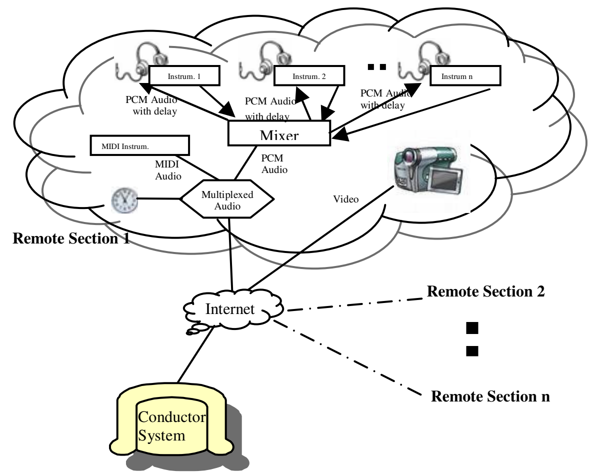


Fig. 5.1 The Distributed System Experimental Prototype

The MIDI stream generated is extended with one special additional event, called the attack event, which is inserted in the MIDI event stream at the beginning of the performance just before the first note. The presence of this event will explicitly determine the start of the performance.

The TANDEM language synchronizes multiple streams based on the synch points identified by the packets. For a reliable performance the streams are buffered at arrival. Due to varying tempos both within sections of the orchestra and between sections, there is no guarantee that a beat will correspond exactly with a synch point, the synchronization is actually performed at the nearest synch point to the beat, or to the time indicated by a particular gesture. In order to meet synchronization needs in the orchestral domain, the synch points will be chosen so that any variation from the beat or action time is below the perceptual level.

The signals which make up the streams contain data which includes both audio PCM data and data related to the score - either MIDI-type messages or simple beat-based information. The data also contains implicit time stamps related to synch points. For live streamed data from multiple remote sites, an atomic clock or similar mechanism may be used to provide a precise enough time-stamp that the combined performance is close enough to perfect synchronization to be under the perceptual level.

It is sometimes impossible to deliver remote performance data in time to avoid perceptual distortion. This may be caused by transient high network latencies. In this case, the Active Repository causes the delayed stream to be muted, rather than allowing the distortion to affect the performance of the orchestra as a whole. Once the stream has caught back up, it will be restarted. This will result in some of the data for the late arriving stream being skipped.

7. RELATED RESEARCH

In this paper we provided a model for the performance of a distributed orchestra. The architecture of the conducting system takes advantage of the synchronization abilities of the TANDEM environment via triggers and the Active Repository, providing an effective way to synchronize live media streams. For this purpose TANDEM has been extended with constructs to describe the conductor's gestures and the semantics of those constructs has been provided. The PCM live streams with at least one MIDI stream per section are multiplexed at each remote source and time stamped before the transmission. The inclusion of MIDI data allows for recognition of beats in the stream for synchronization.

There have been some related efforts in distributed musical performance, however most existing systems that are not sequencers (i.e. software or hardware to create and manage computer generated music) use prerecorded MIDI instruments or MIDI files only. For example, the virtual conducting system described in [4] uses prerecorded MIDI files played locally. More interesting is the approach presented in [23] where an architecture for the management of a distributed musical performance is given. The system does not use a conductor, the stream management again uses only MIDI sequences. In [12] and [22] one-way streaming of musical rehearsal using real time PCM audio was used but all players, including a human conductor, were at a sender site with performance at the receiver.

In Gu [11] PCM audio was streamed over the network in real time in compressed format. To perform compression at a realistic time only prerecorded audio was streamed instead of live performance. The focus of the work was related to compression scheme, packet loss and quality of the streamed audio.

REFERENCES

[1] G. Berry, G. Gonthier, "The ESTEREL Synchronous Programming Language: Design, Semantics, Implementation", *Sci. of Comp. Progr.* **19**, n. 2, pp.87-152, Nov. 1992.

[2] G. Berry, "The Foundations of Esterel", in *Proof, Language and Interaction: Essays in Honour of Robin Milner*, G. Plotkin, et al. ed., MIT Press, pp.425-454, June 2000.

[3] N. D. Binh, E. Shuichi and T. Ejima, "Real-Time Hand Tracking and Gesture Recognition System", *ICGST Int. J. on Graphics, Vision and Image Processing*, **7**, pp.39-45, 2007.

[4] J. Borchers, E. Lee, W. Samminger, M. Mühläuser, "Personal Orchestra: A Real-Time Audio/Video System For Interactive Conducting", *Mult. Syst.,* **9**, pp.458-465, Springer, 2004.

[5] A. Guercio, A. Bansal, T. Arndt, "Languages Constructs and Synchronization in Reactive Multimedia Systems", *ISAST Trans. on Comp. and Soft. Eng.,* **1**, n.1, pp.52-58, 2007.

[6] A. Guercio, A. K. Bansal, "Towards a Formal Semantics for Distributed Multimedia Computing", *Proc. of DMS 2007*, San Francisco Sept. 6-8, pp.81-86, 2007.

[7] A. Guercio, A.K. Bansal, T. Arndt, "Synchronization for Multimedia Languages in Distributed Systems", *Proc. of DMS 2005*, Banff, Canada, Sept. 5-7, pp.34-39, 2005.

[8] A. Guercio, A. K. Bansal, "TANDEM – Transmitting Asynchronous Nondeterministic and Deterministic Events in Multimedia Systems over the Internet", *Proc. of DMS 2004,* San Francisco, pp. 57-62, Sept. 2004.

[9] A. Guercio, A. K. Bansal, "A Model for Integrating Deterministic and Asynchronous Events in Reactive Multimedia Internet Based Languages", *Proc. of the 5th Int. Conf. on Internet Computing (IC 2004)*, Las Vegas, June 21-24, pp.46-52, 2004.

[10] R. Gordon, S. Talley, *Essential JMF – Java Media Framework*, Prentice Hall, 1999.

[11] X.Gu, M. Dick, Z.Kurtisi, U. Noyer, L. Wolf, "Network-centric music performance: Practice and Experiments", *IEEE Comm. Mag.,* **43**, n.6, pp.86-93, 2005.

[12] D. Konstantas, "Overview of telepresence environment for sitributed musical rehersal", *Proc. of ACM Symposium on Applied Computing (SAC'09),* Atlanta, 1998.

[13] G. Luck, J.A. Sloboda, "An investigation of Musicians' Synchronization with Traditional Conducting Beat Patterns", *Music Perform. Res.,***1**,1, pp.6-46, IISN-7155-9219, 2007.

[14] G. Luck, S. Nte, "An Investigation Of Conductors' Temporal Gestures And Conductor-Musician Synchronization, And A First Experiment", *Psychol. of Music,* **36**(1), pp.81-99 2008.

[15] L.A. Ludovico, "Key Concepts of the IEEE 1599 Standard", *Proc. of the IEEE CS Conf. The Use of Symbols To Represent Music And Multimedia Objects*, pp.15-26, Lugano, CH, 2008.

[16] B. McElheran, "*Conducting Technique for Beginners and Professionals*", Oxford University Press, USA, 1989.

[17] Object Manag. Group, "Control and management of A/V streams specification", OMG Doc. telecom/97-05-07, 1997.

[18] M. Rudolf, "*The Grammar Of Conducting"*, Wadsworth, London, 1995.

[19] SMIL2.0 Specification, http://www.w3.org/TR/smil20/ 2001.

[20] G. Stetten, et al., "Fingersight: Fingertip visual haptic sensing and control", *Proc. of IEEE Int. Workshop on Haptic Audio Visual Env. and their Appl.*, pp.80-83, 2007.

[21] XML and Music", http://xml.coverpages.org/xmlMusic.html

[22] A. Xu, et al., "Real-time Streaming of Multichannel Audio Data Over the Internet" , *J. Audio Eng. Soc,* **48**, pp.7-8, 2000.

[23] R. Zimmerman, E. Chew, S Arslan Ay, M Pawar, "Distributive Musical Performances: Architecture and Stream Management", *ACM Trans. on Mult. Comp., Comm. and Appl.*, **4**, n. 2, Article 14, May 2008.

Semantic Composition of Web Services

Manuel Bernal Llinares, Antonio Ruiz Martínez, Mª Antonia Martínez Carreras, Antonio F. Gómez Skarmeta
Department of Information and Communication Engineering
Faculty of Computer Science
University of Murcia
Murcia, Spain
{manuelbl, arm, amart, skarmeta}@um.es

Abstract—**Nowadays the number of applications and processes based on Web Services is growing really fast. More complex processes can be achieved easily through the composition of Web Services. There are proposals like WS-BPEL to compose Web Services but nowadays this process is done statically. There is a strong coupling between the Web Services that are involved in the composition and the composition process itself, thus, changes on the services will invalidate the composition process. To resolve this problem we have defined an architecture where the composition processes are abstract and semantic information is used for linking them to the right Web Services for every situation.**

Collaborative Environments, business processes, service composition, semantic.

I. Introduction

Service-Oriented Architecture (SOA) is the platform for under the Web services technology which has demonstrated to fit with ithaving all the required components defined in SOA: a way to describe services, including the basic information defined in SOA and some more: Web Service Definition Language (WSDL)[3]; a mechanism to represent the necessary messages: SOAP[36]; a service to be able to know the existence of services, a mechanism to search for a services: Universal Description, Discovery and Integration (UDDI)[4].

But the related standards of Web services go far away from the basis of SOA. We also can find: Web Services Interoperability (WS-I)[4]; Web Services Business Process Execution Language (WS-BPEL), an orchestration language using Web services; Web Services Choreography Definition Language (WS-CDL)[5], a choreography language for Web Services; Web Services Choreography Interface (WSCI)[7], a language for describing interfaces used to specify the flow of messages at interacting Web Services.

Web services technology has become the favorite platform over which companies and institutions implement all their services, this heterogeneity of Web services providers and consumers has motivated an increased interest for the composition of services in the research community. This key area of Web services is where the work presented in this paper has been developed. More precisely, the aim of this paper is depict the building of an architecture for composing services according to an abstract description of the process and the use of semantic for annotating services.

The remainder of the paper is organized as follows. We first give some related work about the problem of Web services composition in section 2, then we introduce a motivating scenario in section 3, next we present our solution in section 4. Finally, we give conclusions and future work in section 5.

II. Related Work

Previous work related to Web services composition have taken approaches from the semi-automatic composition[8,9] where a system is built to aid the user in the process of composing Web services (using semantic information to filter the available services and presenting only those that are relevant); to the automatic composition of Web services where the work is mostly focused on the view of the service composition as a planning problem; thus the process is done through the use of HTN[10,11,24], Golog[12-14], theorem proving [15-18], ruled based planning[19,20], model checking[21-23], Case Based Reasoning (CBR)[25], Propositional Dynamic Logic based systems[26], classic AI planning[27], etc.

The composition of services presents two main challenges, one of them related to the orchestration of the services and the other one related to the heterogeneity of the data. Although all the solutions address the problem of the orchestration of services, either aiding the user in the manual composition of services (by filtering information) or defining complex semantic structures (with preconditions and post-conditions that characterize the goal that must be achieve by the orchestration of services, and then using some of the mentioned approaches to automatically create the composition process) very few address the problem of data heterogeneity.

There are really few proposals that give support to the industry standard for the composition of services (WSBPEL) in an automated way, the work presented in this paper fills this gap using semantic information.

III. Motivating Scenario

WSBPEL is an XML-based process/workflow definition execution language, it defines a model and a grammar for describing the behavior of a business process based on interactions between a process and its partners, these interactions occur through the Web service interface of each partner.

WSBPEL shows itself not flexible at all with the underlying services it is orchestrating, changes on those services will affect the orchestration defined in WSBPEL making it unusable. Thus, there is a strong coupling between the business process and the Web services it orchestrates.

Our work is focused on removing this coupling using semantic information. The main advantage of our solution is that it brings adaptability and fault tolerance to the industry standard in the composition of services, providing some grade of portability of business processes from one system to another.

IV. Decoupling Business Processes from Underlying Services

WSBPEL is defined by two XML Schemas[2]:

- Abstract: an abstract process is a partially specified process, it is not intended to be executed as it is. This type of process may hide some information of the required concrete operational.
- Executable: an executable process is fully specified and therefore it can be executed.

To decouple the business process from the underlying services that are involved in it, the abstract definition of WSBPEL is going to be used and the work will focus on how to transform the abstract definition of a composition of services into an executable one.

A WSBPEL document (which describes an orchestration of Web services) is a sequence of steps where some of them involve an operation of a Web service as can be seen (marked with a red circle) on the following figure.

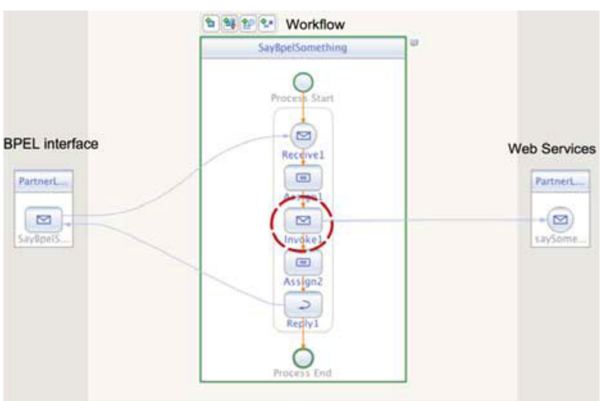


Figure 1. Simple BPEL diagram.

The WSBPEL relays on the WSDL description of Web services to orchestrate them, but this information guarantees only the syntactic interoperability among Web services and, in several cases, this is not enough to ensure that a business process is correctly assembled. Ideally, a business process definition should describe the orchestration in terms of the kind of Web services involved, rather than specifying concrete Web services.

The abstract definition of a business process keeps the workflow but removes all the links to the Web services involved, making the business process independent of the underlying services but unusable as it is. There is no possible way to restore the original business process by hand, and to accomplish it automatically, additional information is needed both on the Web services description and the WSBPEL document.

This additional information is introduced both on the services and the business process by extending its definitions (WSDL and WSBPEL) with SAWSDL[28] annotations, which reference concepts in an ontology. The main advantage of SAWSDL is that it is independent of the ontology language used, thus it is possible to use different formalisms according to the needs of a particular domain.

On the side of the Web services, these SAWSDL annotations[11] defines how to add semantic information to describe several parts of the WSDL document such as input and output messages structures, interfaces and operations. In this work the attribute “modelReference” will be used on the operations of the Web services to describe, semantically, which is the goal they are able to achieve.

On the side of the WSBPEL document, this attribute will be used on every step an operation from a Web service is involved to specify the goal the operation is required to accomplish.

At this point, we have annotated Web services and an annotated abstract business process that we need to translate into an executable one before it can be usable. The translation of the abstract process is done looking for the Web services suitable to accomplish the goal required in each step where a service is involved, thus the annotated available Web services must be reachable somewhere where they can be searched given a goal.

For this, we have developed a Composition Engine that is one of the main components in the architecture of the Semantic System shown in the ECOSPACE[30,31] project.

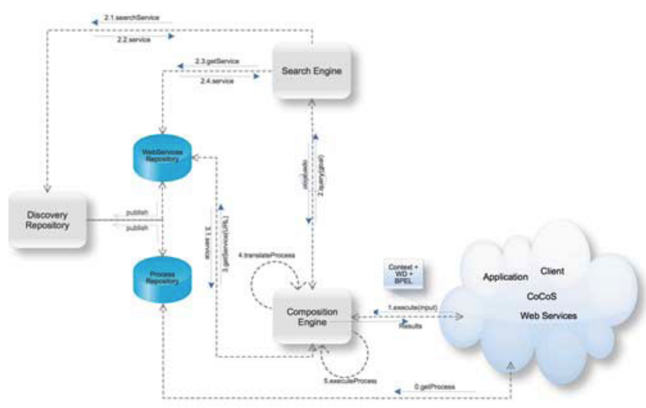


Figure 2. Semantic System Architecture

The Composition Engine interacts with two main components of this architecture.

The Discovery Repository is the component responsible for storing the annotated Web Services descriptions and related artifacts, e.g. SPARQL based pre and post-conditions.

We would like to make a logical distinction between a registry and repository to eliminate any confusion. The term "registry" in its implementation refers to a metadata store; it is analogous to a books catalogue which can be found in a library. The term "repository" refers to the actual content that needs to be stored in addition to its metadata. A repository is analogous to actually book shelf in a library that stores all the books.

The registry and repository infrastructure represents mainly three registry and repositories i.e. Service Registry, Service repository and Ontology registry. Other applications and architectural components (such as Semantic Service Discovery Engine which will be detailed later) can locate the required resources (i.e. service descriptions and ontologies) through registry and repository infrastructure. Detailed discussion about the Service Registry and Repository specification can be found in the deliverable D3.2 of the ECOSPACE project[32].

The Dynamic Semantic Service Discovery (DSSD) is a software component that implements dynamic discovery of Web services, taking into account the preconditions and post-conditions defined in their SAWSDL descriptions. The DSSD comprehends two main subcomponents: the Semantic Registry and the Discovery Agent.

The Semantic Registry of the DSSD acts as an internal library of Web Services operations, and maintain specific data structures holding the definition of preconditions, and the descriptions of post-conditions. The Semantic Registry is coupled with a traditional Registry, that in ECOSPACE architecture is implemented by the Discovery Repository, which holds the SAWSDL descriptions. The Semantic Registry of the DSSD fetches SAWSDL descriptions from the Discovery Repository, and process them in order to extract the semantic information linked by the URIs in the "modelReference" attributes (as describe above). The Semantic Registry uses such each SPARQL CONSTRUCT queries to build the RDF graph corresponding to the effects of the Web Service operation, and store preconditions.

The Discovery Agent is a specialized software component; it has a knowledge base (i.e. a formal description of some information that is known to the agent), ant it accepts a goal (i.e. the description of an objective). The Discovery Agent searches a Web Service operation whose effects allows for the achievement of the goal. The Discovery Agent interacts with the Semantic Registry in order to explore the effects of Web Services operations, and to verify the satisfiability of their preconditions using the information contained in its knowledge base.

Further description is available in[33].

The Composition Engine has a web service interface and offers the execution of an abstract business process (annotated semantically) as if it were executable in a completely transparent way.

Context information like a world description and an invocation context is provided, as well as the abstract WSBPEL document that describes the business process to execute.

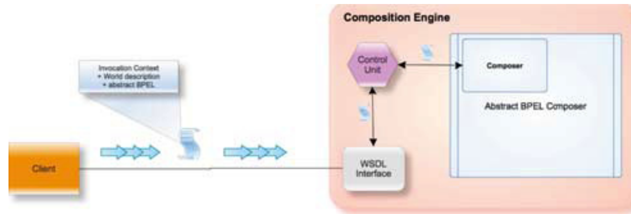


Figure 3. Interface and Control Unit

The translation and execution of the annotated abstract business process is carried out trough several steps.

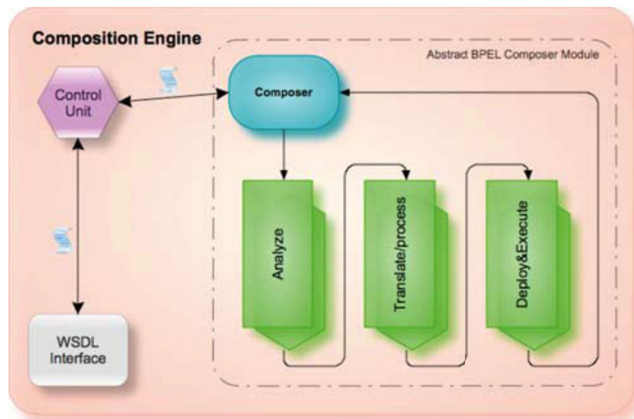


Figure 4. Composition Engine, process overview

In the figure above these lines there is an overview of the translation process. The "Composer" will be responsible of driving the whole process dynamically adapting the behavior of the Composition Engine depending on the context information.

The first stage of the translation is the analysis of the annotated abstract business process. All the goals referenced in the WSBPEL are extracted and used with the context information to query the DSSD for suitable Web services. The DSSD will use that information to look up in the registry where the services are published.

The information collected in this stage is a list where every goal is paired with the most suitable Web service that is able to achieve it.

On the next stage that information is used to translate the WSBPEL into an executable business process. This is the most complex stage because of the flexibility of the WSBPEL language, here, the descriptions of the Web services selected are adapted to meet the syntactical requirements of the WSBPEL in case they don't meet them.

The last stage is where the executable process, obtained before, is deployed in a BPEL Engine like ActiveBPEL[34] or Glassfish[35] (which are the two BPEL Engines considered for this development). Then, the business process is executed and the Composition Engines returns the results back to the client.

This newly created business process must be undeployed from the BPEL Engine after its execution because it is not intended to have a lifespan beyond the execution requested to the semantic system.

V. CONCLUSIONS AND FUTURE WORK

This paper presents an important contribution to solve a key issue of the Composition of Services from the industry promoted standard: the highly coupling between the business process and the underlying services.

This work introduces adaptability to the composition of services not only taking into account possible changes on them, but also introduces the ability to select the most suitable services depending on the context the business process is being execute (i.e. based on costs, requirements, prohibitions, user preferences…). It is providing context-awareness[37] to the composition of services.

At this time the two first stages of the Composition Engine are completed and the executable process result of the translation at the second stage has been proved to work on Glassfish. Our future work includes the implementation of the last stage with the difficulty that every BPEL Engine has its own custom artifacts that need to be created around the WSBPEL in order to deploy the process and there is neither API nor automatic way to do it programmatically, so a lot of effort must be done to implement the last stage.

REFERENCES

[1] Abhijit Patil, Swapna Oundhakar, Amit Sheth, and Kunal Verma. “Meteor-s web service annotation framework”. In Proceedings of the 13th International World Wide Web Conference, New York, USA, May 2004.

[2] OASIS, Web Services Business Process Execution Language Version 2.0, wsbpel-v2.0.pdf.

[3] W3C. Web Services Description Language (WSDL). Online: http://www.w3.org/TR/wsdl

[4] T. Bellwood et al. Universal Description, Discovery and Integration specification (UDDI) 3.0. Online: http://uddi.org/pubs/uddi-v3.00-published-20020719.htm

[5] WS-I, Web Services Interoperability Organization. Online: http://www.ws-i.org/Docs/Brochures/WS-I%20Overview.pdf

[6] W3C, Web Services Choreography Description Language (WS-CDL). Online: http://www.w3.org/TR/2004/WD-ws-cdl-10-20041217/

[7] W3C, Web Services Choreography Interface (WSCI). Online: http://www.w3.org/TR/wsci/

[8] Evren Sirin, James Hendler and Bijan Parsia. “Semi-automatic Composition of Web Services using Semantic Descriptions”.

[9] David Trastour, Claudio Bartolini and Javier Gonzalez-Castillo. “A Semantic Web Approach to Service Description for Matchmaking of Services”

[10] Sirin E., et al., HTN Planning for Web Service Composition Using SHOP2. Web Semantics Journal. 2004. 1(4): p. 377-396.

[11] Sirin E., B. Parsia and J. Hendler. Template based composition of semantic web services, in AAAI fall symp on agents and the semantic web. 2005: Virginia, USA.

[12] Narayanan, S. and S.A. McIlraith. Stemulation, verification and automated composition of Web services. In The 11th International World Wide Web Conference. 2002. Honolulu, Hawaii, USA.

[13] McIlraith, S.A., T.C. Son, and H. Zeng, Semantic Web Services. IEEE Intelligent Systems, 2001. 16(2): p. 46-53.

[14] McIlraith, S. and T.C. Son. Adapting Golog for composition of Semantic Web services. In Knowledge Representation and Reasoning (KR2002). 2002. Toulouse, France.

[15] Waldinger, R.J., Web Agents Cooperating Deductively, in Proceedings of the First International Workshop on Formal Approaches to Agent-Based Systems-Revised Papers. 2001, Springer-Verlag.

[16] Lämmermann, S., Runtime Service Composition via Logic-Based Program Synthesis, in Department of Microelectronics and Information Technology. 2002. Royal Institute of Technology.

[17] Rao, J., P. Kungas, and M. Matskin. Application of Linear Logic to Web Service Composition, in The 1st Intl. Conf. on Web Services. 2003.

[18] Rao, J., P. Kungas and M. Matskin. Logic-based Web services composition: from service description to process model, in The 2004 Intl Conf on Web Services. 2004. San Diego, USA.

[19] Ponnekanti, S.R. and A. Fox, SWORD: A Developer Toolkit for Web Service Composition, in The 11th World Wide Web Conference 2002: Honolulu, Hawaii. USA.

[20] Medjahed, B., A. Bouguettaya, and A.K. Elmagarmid. Composing Web services on the Semantic Web. VLDB Journal. 2003. 12(4).

[21] Kuter, U., et al. A Hierarchical Task-Network Planner based on Sybolic Model Checking, in The International Conference on Automated Planning & Scheduling (ICAPS). 2005.

[22] Traverso, P. and M. Pistore. Automated Composition of Semantic Web Services into Executable Processes, in The 3rd International Semantic Web Conference (ISWC2004). 2004.

[23] Pistore, M., et al. Automated Synthesis of Composite BPEL4WS Web Services, in IEEE Intl Conference on Web Services (ICWS’05).

[24] Massimo Paolucci, Katia Sycara and Takahiro Kawamura. Delivering Semantic Web Services. WWW2003.

[25] Benchaphon Limthanmarhon and Yanchun Zhang. Web Service Composition with Case-Based Reasoning, in 14th Australian Database Conference 2003. Adelaide, Australia. Conferences in Research and Practice in Information Technology, Vol. 17.

[26] Daniela Berardi, Diego Calvanese, Giuseppe De Giacomo, Richard Hull and Massimo Mecella. Automatic Composition of Transition-based Semantic Web Services with Messaging, in Proceedings of the 31st VLDB Conference, Trondheim, Norway, 2005.

[27] Rao, J., et al., A Mixed Initiative Approach to Semantic Web Service Discovery and Composition: SAP’s Guided Procedures Framework, in The IEEE Intl Conf on Web Services (ICWS’06). 2006.

[28] Semantic Annotations for WSDL Working Group website, http://www.w3.org/2002/ws/sawsdl/

[29] Kopecký, J.; Vitvar, T.; Bournez, C. and Farrell, J. (2007) SAWSDL: Semantic Annotations for WSDL and XML Schema. IEEE Internet Computing, 2007, 11, 60-67.

[30] ECOSPACE project, http://www.ip-ecospace.org/

[31] ECOSPACE Deliverable 3.7 “Final version of the augmented middleware”.

[32] ECOSPACE Deliverable 3.2 “Middleware Open Interfaces and Service Support Prototype”.

[33] Iqbal, K.; Sbodio, M. L.; Peristeras, V. and Guiliani, G. (2008) Semantic Service Discovery using SAWSDL and SPARQL, Proceedings of the SKG 2008, IEEE Press, 2008, (yet to appear)Bowman, M., Debray, S. K., and Peterson, L. L. 1993. Reasoning about naming systems. ACM Trans.

[34] ActiveEndpoint, The ActiveBpel Community Edition BPEL Engine. http://www.activevos.com/community-open-source.php

[35] Glassfish Community. https://glassfish.dev.java.net/

[36] W3C, SOAP specification. http://www.w3.org/TR/soap/

[37] Anind K. Dey, Gregory D. Abowd. Towards a Better Understanding of Context and Context-Awareness. GVU Technical Report GIT-GVU-99-22, College of Computing, Georgia Institute of Technology, 1999.

Proceedings
International Workshop on
Distant Education Technology
(DET 2009)

Co-Editors

Paolo Maresca, *University Federico II, Napoli, Italy*
Qun Jin, *Waseda University, Japan*

Eclipse: a new way to Mashup.

Paolo Maresca
Dip. Informatica e Sistemistica
Università di Napoli Federico II, Italy
paolo.maresca@unina.it

Giuseppe Marco Scarfogliero
Università di Napoli Federico II, Italy
g.scarfogliero@studenti.unina.it

Lidia Stanganelli
DIST - University of Genoa, Italy
lidia.stanganelli@poste.it

Abstract

In our approach for designing enterprise solutions, there is the need to realize some situational applications to manage all the enterprise business processes that the major Enterprise Applications cannot treat due to the particularities of these processes. The specific nature of these processes and their less relevance in the global mission make them less attractive for software houses and customers due to the high costs of designing and development. So the need to find a solution which guarantees low costs and short times of production. Our interest resides in mashup applications and the web 2.0 capabilities. Our intent, with this paper, is to show the eclipse platform as a very good solution to the problem of designing e developing mashup applications, showing which are the classical levels of a mashup application and how eclipse platform can satisfy all mashup's needs thanks to its modular and flexible structure. In conclusion we show also the aim the governs the eclipse community and the constant rejuvenation process that gives us more trust on the future possibility in this way.

1 Introduction

Today, the Information Technology scenario is having a deep evolution, under the unceasing pressure of Market, that every days shows new needs. This change is led by technology evolution process, which offers innovative business opportunities due to new discoveries.

The Software production sector for enterprises is certainly one of the most interested scenarios by this changing: next to the Enterprise Applications, developed by IT as solution to the largest part of an enterprise business problems, there is the need for Situational Applications, software built ad hoc to manage particular business processes linked to the different realities. Very often the resources destined to the production of these applications are limited, because of the lower relevance that they have in the global mission. The tendency is to adopt low quality software or to use non conventional alternatives, using software built for other purposes to achieve own goals.

The main difficulty to invest in the production of software of this kind is in the "artistic" and "social" nature of the business processes to model, in the sense that their particularity and specificity do not allow their implementation in Enterprise Applications.

So, the challenge is to provide very flexible, agile and low cost methods and processes to develop Situational Applications, in order to exploit the business opportunity represented by the "Long Tail".

The possibilities offered by web 2.0 technologies are some of the most accredited solutions to this problem. In this scenario Mashups have a great relevance.

2 Mashup

A mashup is a lightweight web application, which allows users to remix informations and functions belonging to different sources and to work with them to build software in a completely new, simple and quick way. The user can efficiently model their own business process under the own vision of the problem, achieving a result so particular and specific that is impossible to obtain with the older technologies.

Mashups stands on the fundamental concept of data and services integration; to operate in this way there are three main primitives: Combination, Aggregation and Visualizzation.

The first allows to collect data from heterogeneous sources and to use them within the same application; the second primitive allows to operate on collected data having a measure and building new information starting from them; the last is used to integrate data in a visual way using maps or other multimedia objects.

In a technological view of the Mashup and of its data and services integration problem, the natural representation of the problem itself can be obtained using a level/pyramidal approach.

Fig.01 – The Mashup Pyramid

In the lowest abstraction layer there are Data Feeds and web technologies involved by them. They represent a good solution to access to updated data in a quick and secure way.
In the immediately superior level live the APIs, used to obtain data dynamically and on demand services.
A great level of abstraction is achieved by Code Libraries, that can be thought as application frameworks and API packages built to resolve some kinds of problems.
On the Code Library level stands the Gui Tools level, made of widgets and technologies related to the composition of small graphical applications to show data or to allow the access to a service.
On the top of the pyramid there's the "Platform" level, composed by all the tools and platforms that support mashup applications building, allowing to compose single graphical elements and lower level data.

3 Eclipse

At this point it might be clear to the reader the complexity of this model and the need to act on each of pyramid levels in the application building process in order to obtain a flexible and complete development process of a mashup application. From here the need of an integrated development environment, capable to adapt itself to each kind of need thanks to its modular and flexible architecture, allowing to face every aspect of the mashup problem and that drives the developer through all the production process till deployment and testing of the final application.
These requisites are well satisfied by the Open Source Development Platform "Eclipse", that can greatly adapt itself to every scenario thanks to its modular architecture. Integrability is one of the main directives of the Eclipse project from its birth: the platform architecture allows 5 different integration levels as represented in the following diagram.

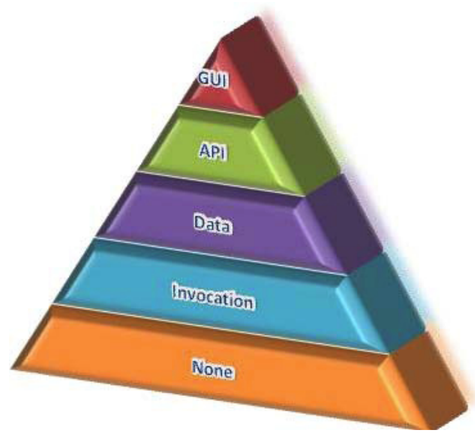


Fig. 02 – The Eclipse Integration Pyramid

The lowest level is "None – No Integration" which represents the possibility to have no integration with other external tools if this integration is not needed.
The "Invocation" level represents the integration obtained by invocation of tools and services external to Eclipse within the platform itself. Services are executed as external processes distinct from the IDE one, using the same eclipse resource manager to start them. Platform gives the possibility to manage a tool-resource association registry independent from the Operative System one.
"Data" level is certainly the one that offers the greatest level of integration. Eclipse platform, in fact, allows to collect data from heterogeneous sources, to give them a structure and to provide them to own applications, in a coherent and very flexible way.
The "API" integration level graft perfectly on Data level. The extreme flexibility of the Data level is balanced by the need of decode, understand and maintain integrity of Data. Using APIs allow to access data in a coherent, secure and especially dynamic way, so the programmer can release the burden of dealing of the explicit manage of data. With APIs there is the introduction of the concept of service, intended as an on demand action on data. The modular structure of eclipse allows each application to define its own APIs and services that become usable by the platform itself and by its components.
On the top of the pyramid there is GUI integration level, which allows many tools or application to share the platform Graphical User Interface becoming an unique application perfectly integrated in the IDE structure, starting from different applications.

4 Points of Convergence

There is a clear correspondence between the Mashup pyramid levels and the Eclipse Integration Piramyd ones:

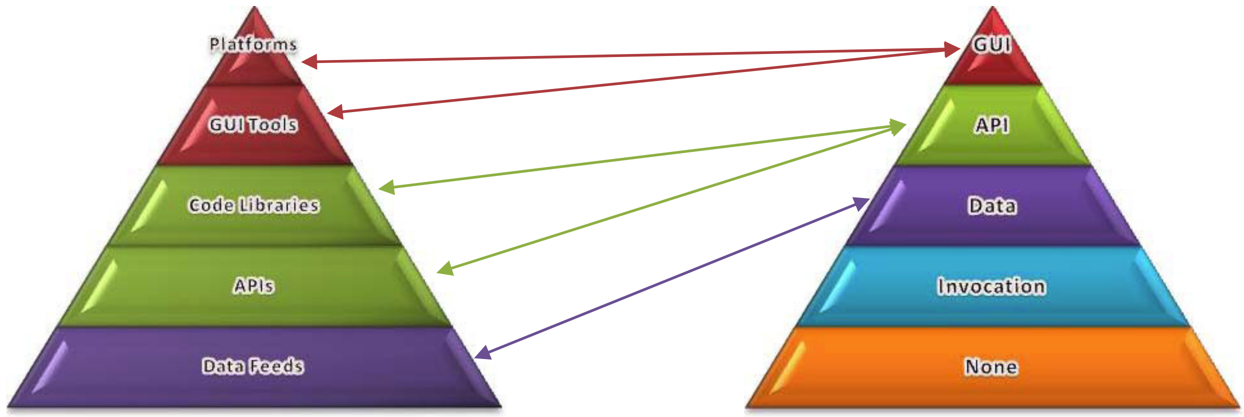


Fig. 03 – Corrispondence between pyramids

The "Data" level of Eclipse Integration Pyramid allows to greatly manage Data Feeds, base of mashup pyramid, extending this possibilities to all other structured Data belonging to other sources like heterogeneous Databases. This perspective appear very interesting in building enterprise mashups, that realize the convergence between data belonging to Enterprise Databases and data belonging to web services external to own enterprise infrastructure.
Eclipse Galileo will offer many opportunities in this way, including Data Tool Platform (DTP) project and the famous Business Intelligence and Reporting Tool (BIRT), that allow to collect and structure data using the Open Data Access (ODA) framework, which realizes the connection with the most common data sources: XML, Web Services, CSV files and JDBC. Obtained data can be easily managed by the known middle level eclipse framewoks and be the base for EMF applications or others.
The "API" level allows to realize integration through platform API and Plugins that compose the particular installation. The modular structure of eclipse makes easy to use external APIs or Code Libraries in a native manner or managing them through particular plugin. A famous example of the last possibility is offered by the "gdata-java-client-eclipse-plugin" which, after installed, gives the opportunity to easily create Java application that uses the common Goolgle APIs. These possibilities make the platform itself a natural candidate in realizing the right integration required from Mashup's "API" and "Code Libraries" levels.
The "GUI" level is certainly one of the most powerful and tested integration level in eclipse. The extreme simplicity that characterizes the extension of the development environment and its graphical personalization makes the platform adapt to realize any kind of application, beginning from different applications too, using perspectives, views and editors. So, eclipse results to be the perfect environment in which integrate mashup application widgets directly in its architecture, with the whole flexibility, support and easiness of integration in the platform. Eclipse is a unique environment in which realize the development of the environment itself.
Last fundamental step is to bring the realized eclipse mashup application on the web. Because of its genesis as stand-alone software development tool, sometimes are not clear the real possibility of eclipse in the web 2.0 field. There are many projects that allow the platform to be accessible and usable from the web using a common browser. Among all these projects one of the most interesting is the "Eclifox" plugin developed as IBM Alphawork; it makes available a remote eclipse instance on the web through Jetty web server, transforming SWT based GUIs on XUL based GUIs. XUL is the famous language used by Mozilla products like Firefox. Another important perspective is brought by the project "Rich Ajax Platform" (RAP), that will be a component of Eclipse Galileo having the maximum compatibility with the platform. This Project allows to design Ajax applications based on eclipse in a simple way very similar to RCP Application building, substituting SWT widget library with RWT built for web. So RAP is a very good candidate to mashup application's GUI building, because the entire application is transformed in a web 2.0 application, using the common Java technologies for server-side programming without the need for an eclipse instance running on a server.

5 Web Services

One of the most interesting data and services source for mashups is represented by web services, because using them allows to link services belonging to Enterprise SOA and services belonging to an external WOA. Actually service integration in eclipse is managed by Data level through ODA drivers or by API level through specific plugins. A new scenario will be born with Galileo based on eclipse 3.5 that will furnish a major support for web services.
Essentially WOA see figure 4, that is a subset of SOA, describes a core set of Web protocols like HTTP and plain XML as the most dynamic, scalable, and

interoperable Web service approach. The only real difference between traditional SOA and the concept of WOA is that WOA advocates REST, an increasingly popular, powerful, and simple method of leveraging HTTP as a Web service in its own right.
some plain old XML to hold your data and state to top it all off.

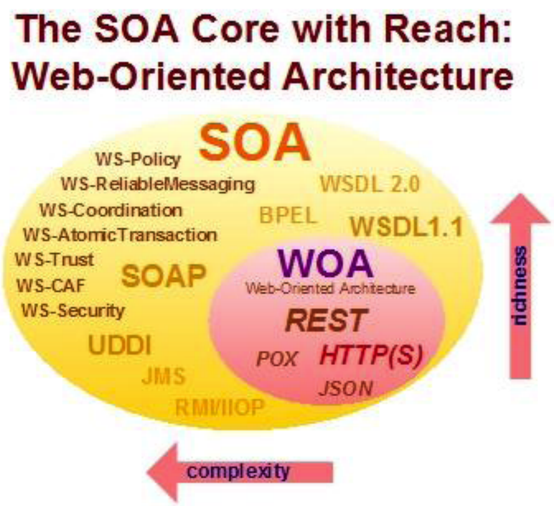


Fig. 04- SOA and WOA comparison architecture

WOA architecture emphasizes generality of interfaces (UIs and APIs) to achieve global network effects through five fundamental generic interface constraints:

1. Identification of resources
2. Manipulation of resources through representations
3. Self-descriptive messages
4. Hypermedia as the engine of application state
5. Application neutralità

This generalization enable us to match easily WOA resources with Mashup Pyramid (see fig. 3).

6 Eclipse and Jazz

Another great advantage in using eclipse is the convergence in act between the eclipse project and Jazz platform: the introduction of Jazz candidates eclipse as a complete tool which allows the collaborative development and the managing of the whole software life cycle. These innovations perfectly agree with mashup philosophy.
Jazz is an IBM initiative to help make software delivery teams more effective, Jazz transform software delivery making it more collaborative, productive and transparent.
The Jazz initiative is composed of three elements:
- An architecture for lifecycle integration
- A portfolio of products designed to put the team first
- A community of stakeholders.
An architecture for lifecycle integration
Jazz products embody an innovative approach to integration based on open, flexible services and Internet architecture. Unlike the monolithic, closed products of the past, Jazz is an open platform designed to support any industry participant who wants to improve the software lifecycle and break down walls between tools.
A portfolio of products designed to put the team first
The Jazz portfolio consists of a common platform and a set of tools that enable all of the members of the extended development team to collaborate more easily. The newest Jazz offerings are:
• Rational Team Concert is a collaborative work environment for developers, architects and project managers with work item, source control, build management, and iteration planning support. It supports any process and includes agile planning templates for Scrum and the Eclipse Way.
• Rational Quality Manager is a web-based test management environment for decision makers and quality professionals. It provides a customizable solution for test planning, workflow control, tracking and reporting capable of quantifying the impact of project decisions on business objectives.
• Rational Requirements Composer is a requirements definition solution that includes visual, easy-to-use elicitation and definition capabilities. Requirements Composer enables the capture and refinement of business needs into unambiguous requirements that drive improved quality, speed, and alignment.
Jazz is not only the traditional software development community of practitioners helping practitioners. It is also customers and community influencing the direction of products through direct, early, and continuous conversation. Fig. 5 shows Db2 on campus project community monitored by using Jazz tools. The project organization of the project was 130 students 4 thesis student about, and was stimulated by using team concert application. Jazz is also a process definition framework including agile and personalized processes.

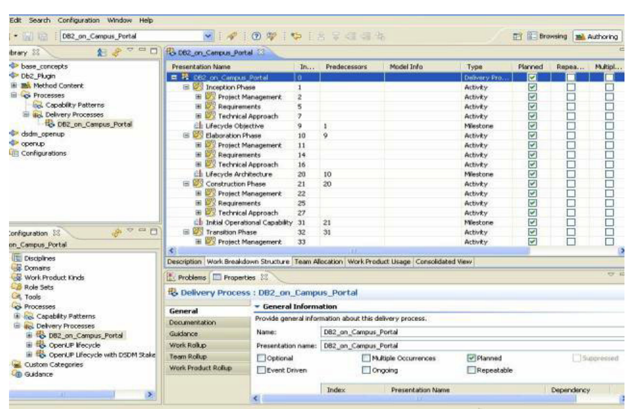


Fig.05 – Db2 on campus project - Jazz

7 CityInformation: a mashup example using BIRT

To underline the real possibilities of eclipse in mashup developing, we show CityInformation, a simple example on how eclipse BIRT project can be used to realize a mix of data belonging to different data sources.

CityInformation shows to the user some information on an user chosen American City in the form of a BIRT HTML report.

When the application starts, it asks the user to insert the name of the city to display information (Fig.06).

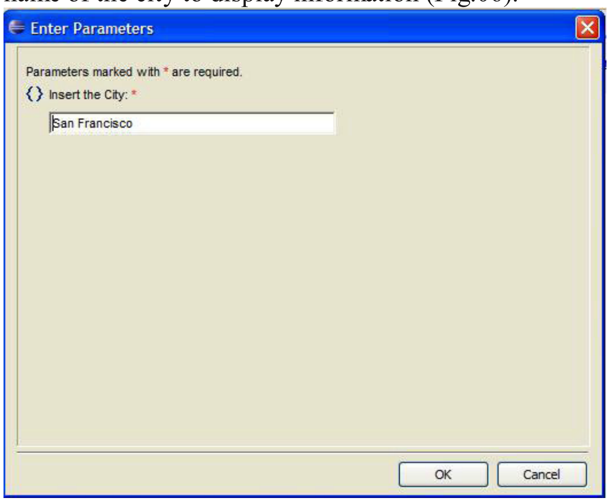


Fig.06 – Enter Parameters

Then the application invokes some free web services to retrieve some information on the city:

The webservice **WeatherForecast** [cfr. Biblio12.] supplies weather forecast information for all the week and the geographic position of the city. Longitude and Latitude are used to display the city map by **Google Maps** using a mashup with an external website. Under the map, forecast information are displayed grouped by day, showing an image and the expected temperatures.

The **Amazon** webservice [cfr. Biblio14.] is used to obtain a list of the most sold Travel Guides of the City on Amazon.com; each book is displayed to the user with its own cover image.

Fig.07 shows the report obtained requesting information on the city of San Francisco.

8 Conclusions and future development

In this paper we showed our belief in mashups as solution to Situational Application development and the need of an integrated environment in which exploit all the possibilities given by mashup philosophy. We believe that eclipse platform is a very good candidate for this purpose thanks to its modular and flexible architecture that allows to manage every abstraction level of the mashup pyramid in a simple way.

As future development we aim at integrating first and second mashup pyramid with the corresponding two

Fig. 07 – City Information Report on San Francisco

eclipse levels. Facing with the next Galileo release of eclipse that could be released by June 2009. A common project is also growing grouping together Napoli and Salerno University with IBM and their business partner with the aims to research new mashup methodologies, technologies and best practices. This collaboration is a great opportunity to integrate knowledge belonging to these different realities, mashing together open-source solutions, university's resources and technologies from enterprises development environments, and to have the possibility to prove that eclipse and mashups can be the base on which build solutions to many problems of modern enterprises and organizations.

Bibliography.

1 Maresca P. (2009)

La comunità eclipse italiana e la ricerca nel web3.0: ruolo, esperienze e primi risultati

Slide show for Mashup meeting at University of Salerno.

2 Raimondo M. (2009)

Web 2.0 Modelli Sociali, Modelli di Business e Tecnologici

IBM Slide show for Mashup meeting at University of Salerno.

3 IBM developerWorks Mashup section

http://www.ibm.com/developerworks/spaces/mashups

4 Duane Merril (2006)

Mashups: The new breed of Web app

http://www.ibm.com/developerworks/web/library/x-mashups.html?S_TACT=105AGX01&S_CMP=LP

5 IBM website

www.ibm.com

6 Jim Amsden

Levels of Integration: five ways you can integrate with eclipse platform

http://www.eclipse.org/articles/Article-Levels-Of-Integration/levels-of-integration.html

7 Eclipse Rich Ajax Platform web site

http://www.eclipse.org/rap/

8 Eclipse Data Tooling Platform web site

http://www.eclipse.org/datatools/

9 Eclipse Business Intelligence and Reporting Tool web site

http://www.eclipse.org/birt/

10 Eclifox web site

http://www.alphaworks.ibm.com/tech/eclifox

11 Weather Forecast webservice: http://www.webservicex.net/WeatherForecast.asmx?WSDL

12 Google Maps API:

http://code.google.com/intl/it-IT/apis/maps/

13 Amazon webservices: http://webservices.amazon.com/AWSECommerceService/AWSECommerceService.wsdl?

Mashup learning and learning communities

Luigi Colazzo, Andrea Molinari
Dip. Informatica e Studi Aziendali
Università di Trento, Italy
luigi.colazzo@unitn.it; andrea.molinari@unitn.it

Paolo Maresca
Dip. Informatica e Sistemistica
Università di Napoli Federico II, Italy
paolo.maresca@unina.it

Lidia Stanganelli
DIST - University of Genoa, Italy
lidia.stanganelli@unige.it

Abstract

The web 2.0, when meeting the virtual communities (VC), creates many issues when communities are closed, but have a great potential if they take advantage of the inheritance mechanism normally implemented in (advanced) virtual communities systems. When a community platform is in place, the system should provide a lot of basic services in order to facilitate the interaction between community's members. However, every community has different needs, every organization that implemements a VC platform needs some special services, every now and then users or organizations request new services.

So, the VC environment is very fertile in terms of personalizations / evolutions / new developments, especially in learning settings. In order to fulfill these growing requests, the developers of e-learning applications have different possibilities: a) build the personalization "from scratch" b) create new web services for the new requests c) using a mashup approach to respond to the requests. In this paper, we will explore the promising perspectives of the latter option. Mashup is an interesting approach to new data / services development, and we will investigate its perspectives in e-learning field. Mashup seems to have a great appealing since it is devoted to the reusing approach that is a typical job in VC ongoing.

1 Introduction

Today, the Information Technology scenario is having a deep evolution, under the uncreasing pressure of Market, that every days shows new needs. This change is led by technology evolution process, which offers innovative business opportunities due to new discoveries.

The Software production sector for enterprises is certainly one of the most interested scenarios by this changing: next to the Enterprise Applications, developed by IT as solution to the largest part of an enterprise business problems, there is the need for Situational Applications, software built ad hoc to manage particular business processes linked to the different realities. Very often the resources destined to the production of these applications are limited, because of the lower relevance that they have in the global mission. The tendency is to adopt low quality software or to use non conventional alternatives, using software built for other purposes to achieve own goals

The main difficulty to invest in the production of software of this kind is in the "artistic" and "social" nature of the business processes to model, in the sense that their particularity and specificity do not allow their implementation in Enterprise Applications.

So, the challenge is to provide very flexible, agile and low cost methods and processes to develop Situational Applications, in order to exploit the business opportunity represented by the "Long Tail" [0].The possibilities offered by web 2.0 technologies are some of the most accredited solutions to this problem. In this scenario Mashups have a great relevance.

A mashup is a lightweight web application, which allows users to remix informations and functions belonging to different sources and to work with them to build software in a completely new, simple and quick way. The users can efficiently model their own business process under the own vision of the problem, achieving a result so particular and specific that is impossible to obtain with the older technologies.

As a initial experiment we would like to discuss following of mashup learning since this could be one of the real case in which we need to adapt the learning requirements to all user needs using a different concept much more relates to web service than the most known functional services offered by learning platform. Learning platform, in several cases, could be viewed as "Hibernate knowledge collection" from which students can learn without adding their own perceptions. In a mashup learning everyone can add his /her personal knowledge by using simple mashup primitives in a mashup learning environment.

Next chapter will cover mashup primitives, Chapter 3 discuss about virtual communities and the mashup tendency growing around such an environment. Chapter 4 will discuss first results and state the conclusions and future development.

2 Mashup and Eclipse

Mashups stands on the fundamental concept of data and services integration; to operate in this way there are three main primitives: Combination, Aggregation and Visualization. First primitive allows to collect data from heterogeneous sources and to use them within the same application; the second primitive allows to operate on collected data having a measure and building new information starting from them; the last is used to integrate data in a visual way using maps or other multimedia objects.

In a technological view of the Mashup and of its data and services integration problem, the natural representation of the problem itself can be obtained using a level/pyramidal approach (see fig. 1).

Fig.1 – The compared Mashup and Eclipse Pyramid

In the lowest abstraction layer there are Data Feeds and web technologies involved by them. They represent a good solution to access to updated data in a quick and secure way.

In the immediately superior level live the APIs, used to obtain data dynamically and on demand services. A great level of abstraction is achieved by Code Libraries, that can be thought as application frameworks and API packages built to resolve some kinds of problems.

On the Code Library level stands the GUI Tools level, made of widgets and technologies related to the composition of small graphical applications to show data or to allow the access to a service.

On the top of the pyramid there's the "Platform" level, composed by all the tools and platforms that support mashup applications building, allowing to compose single graphical elements and lower level data. Model showed in fig. 1 is complex but we have the possibility to operate at each pyramid stage in order to build a flexible and complete process. Obviously we need of a both complete and flexible development process around a stable technology as eclipse.

At this point it might be clear to the reader the complexity of this model and the need to act on each of pyramid levels in the application building process in order to obtain a flexible and complete development process of a mashup application. From here the need of an integrated development environment, capable to adapt itself to each kind of need thanks to its modular and flexible architecture, allowing to face every aspect of the mashup problem and that drives the developer through all the production process till deployment and testing of the final application.

These requisites are well satisfied by the Open Source Development Platform "Eclipse", that can greatly adapt itself to every scenario thanks to its modular architecture. Integrability is one of the main directives of the Eclipse project from its birth: the platform architecture allows 5 different integration levels as represented in the following diagram.

The lowest level is "None – No Integration" which represents the possibility to have no integration with other external tools if this integration is not needed.

The "Invocation" level represents the integration obtained by invocation of tools and services external to Eclipse within the platform itself. Services are executed as external processes distinct from the IDE one, using the same eclipse resource manager to start them. Platform gives the possibility to manage a tool-resource association registry independent from the Operative System one.

"Data" level is certainly the one that offers the greatest level of integration. Eclipse platform, in fact, allows to collect data from heterogeneous sources, to give them a structure and to provide them to own applications, in a coherent and very flexible way.

The "API" integration level graft perfectly on Data level. The extreme flexibility of the Data level is balanced by the need of decode, understand and maintain integrity of Data. Using APIs allow to access data in a coherent, secure and especially dynamic way, so the programmer can release the burden of dealing of the explicit manage of data. With APIs there is the introduction of the concept of service, intended as an on demand action on data. The modular structure of eclipse allows each application to define its own APIs and services that become usable by the platform itself and by its components.

On the top of the pyramid there is GUI integration level, which allows many tools or application to share the platform Graphical User Interface becoming an unique application perfectly integrated in the IDE structure, starting from different applications.

3 Virtual communities

Virtual communities, when applied in organizations (universities, companies, public administrations etc) have a hierarchical structure in nature. This of course is not exactly the typical idea of web 2.0, where contents can be created and aggregated freely by people. A virtual communities system applied to an organization normally requests that the single community is a closed community, where every member has been accepted by the community administrator. This happened also in communities like the ones built in our University, where initially, in the name of free access to everyone, communities were opened. However, after a while, it was clear that the community (mostly associated with the metaphor "course") should be closed only to participants.

This structure of communities related with each other in a hierarchy or in a net is by far more complex than a "flat" architecture, where communities are sort of islands in an archipelago, connected when and if they want. In a virtual communities system like the one developed at the University of Trento, communities are related because they are part of a hierarchy (mostly determined by the organization), but they can be related also trasversally across the different branches of the whole communities tree.

Moreover the philosophy is the one we use in a typical open innovation network of users. In an open innovation group an idea can rise and flow from a community to another one allowing a major self-improvement than a closed community can. An example will clarify the thing. In an academic institution, virtual communities normally can be created simply following the traditional organizational structure of courses, i.e. (in Italian university)

- University – Faculty – Degree – Course ……

This means that we can have the course "DataBase" that is part of the Master Degree in Computer science, that is a community of the community "Faculty of Engineering" that is a sub-community of the "University of Trento" community. This structure has very interesting properties for virtual communities, properties that are typical of any hierarchy: inheritance, propagation, multiple inheritance, polymorphism.

In our virtual communities system, we have another interesting property that is "trasversal inheritance". This means that a community under one branch can inherit data, services or anything else from another community in a different branch. Once again the academic settings have a lot of these examples. Imagine that the above course "Database" of the "Faculty of Engineering" is held by the same teacher also for students of another faculty. In our systems, this means that the students enrolled in the second community should enroll to the first, but that community is in a different branch (Faculty) where normally they do not have access. Trasversality among communities in different branches allow us to create this effect.

On this basis, the mashup ideas exposed above in chapter 1 can offer interesting developments: imagine for example the potentiality of a wiki, developed for the community "database" above, that could be inherited by the trasversal community of the same course held for the students of the other faculty.

Another example is the typical situation where course with high numbers of students are split into different sub-courses, but of course the topics, the material, all the services are shared among the different sub-course. If we have a sub-community "database-a", a subcommunity "database-b", all of them can create an wiki internal to the sub community, but it would be very interesting to aggregate these two wikis into one single wiki set at the level of the parent community. This is a typical problem of mashuping data coming from different communities that have some hierarchical relationship between them.

4. Mashup learning

In e-learning field, the word "mashup" could evoke different perspectives. The first association between the two words probably has been made when the scientific community started to talk about E-learning 2.0. E-learning 2.0 of course is strictly related with web 2.0 metaphor, and the respective ideas of users participation in content production, social networks, blogs, wikis etc. So, in the world of e-learning, the closest thing to a social network is a community of practice, where participarts promoted the interaction and collaboration of people inside the community.

In this environment of cross-fertilization between new web 2.0 tools and e-learning, the basic idea of mashing up services and data finalizing them to educational activities is pretty straightforward. Mashup has also an appeal in terms of authoring environments, where the teacher is able to mash up digital contents originated from different sources.

In general mashup "ideas" and e-learning are in theory highly compatible: we believe therefore that the following argumentations could be accepted as a starting point for further studies:

- Mashup general concept is very interesting and promising in creating / integrating web application
- Mashup enable the open innovation user network collaboration that is a fertile way to flow idea and data from a community to another
- Virtual communities are a particular fertile settings for new services created and available even to the specific and detail need of a single community, and in situational application development
- Some development environments like Eclipse are "philosophically" very close to the central ideas of mashup.
- E-learning settings are closely related with mashup approach in the acquisition and authoring of educational material. Teacher could create new and media-enriched learning objects taking advantage of mashup environments and primitives
- E-learning settings are even more requesting this flexibility in creating / adapting / personalizing services oriented to didactical activities.

So the crucial questions are the following:

- Are there any potential applications for mashup in e-learning / collaboration fields?
- is the current mashup technologies ready for allowing users to create their own mashups in e-learning settings?
- If not, what is missing to mashup philosophy to become a "killer application", or better approach to e-learning development?
- How can authors' rights be identified and protected in mashup-enabled environments?
- On the other side, is it the time to shift from closed innovation user network (web 1.0) to open innovation user network (web 2.0 and 3.0) taking advantage of the metaphor and tools available for virtual communities?
- How will service oriented architecture impact the learning paradigm in the next future ?

Of course there is a general response that could conclude the discussion: mashup is a very interesting and promising approach, all the other difficulties will be overtaken with time and market approval. Mashup editors, like Yahoo Pipes and IBM Lotus Mashup Maker, are available on the market (with different market strategies); they allow end-users, even non-programmer end-users, to mash-up information sources and services to build new information available to satisfy their long-term, or immediate information needs.

Though we agree with this general claim, our first experiments are showing some dark points, and some clarifications that must be done in this area. For example, on the side of tools, with the increasing number of services and information sources available, and with the complex operations that mashup tools tend to stimulate (like filtering and joining), even an easy to use editor is not sufficient [12].

First of all, it must be clarified who is the final user of mashups in learning settings. Here follows a list of possible users of this new paradigm, ordered from the one more involved in technical operations (the programmer) to the less technical user that could mashups some e-learning services (the participant):

- the programmer, that will use enhanced-development environments (like Eclipse) in order to rapid develop mashup services from other services already existing
- the administrator of the e-learning platform, that will assemble some data or services extracted from the e-learning platform based on request
- the teacher, that due to his/her specific didactical needs, is allowed to use some mashup platform (like Yahoo Pipes™) to create new services / data for his/her activities
- the participant, that uses mashup techniques to gather data or services for his/her educational needs

As you see, the panorama is very variegated, with different level of involvement, technical complexity, final objectives. In the case study we are using to understand and deepen this topic, i.e. our Virtual community platform, of course our first problem regards the programmer.

One great advantage of mashuping on these VC systems regards, as mentioned, the possibility of creating very quickly new services for the final users just approaching with a mashup-enabled development platform. This in some way resembles the times of “software reuse”, “software as a component”, and for some reason is contiguous to concepts like SaaS approaches. The difference, anyway, are mainly in the tools, in the general approach to the construction of new services, in the technicalities that allow a mashup-enabled platform to be efficient for programmers.

Regarding the last possible end-user of mashup e-learning, i.e., non-technical users like teachers of participants, of course this is at the moment more a dream than a concrete perspective. What we would like to stress is the potential of this approach. Imagine for example, the general idea of mashup applied to educational material construction, or in the creation of didactical paths that the participants can build with an easy-to-use approach where the contents are aggregated (graphically?) from different, web-based content resourses.

The idea showed here could be strictly connected with the process (or didactical path) that substain the material. We mean when mashuping resources we could mashup also the process that substain them. We need to have Process Re-Engineering Process (PREP) as another mashup experimentation area with their primitives methodologies an tools. Since, mashup includes both processes and products it implies, third area, new distributed architecture systems as peer-to per or service oriented and more prototyping tools as eclipse platform and cooperation-collaboration, as jazz

5. Conclusions and future development

In this paper we showed our belief in mashups as a promising and new approach for e-learning settings, specifically those that are more oriented to create a collaborative environment, like Virtual Communities. Mashup applications/environments/tools are interesting from many different perspectives, from the perspective of the producers of contents (teachers, or in web 2.0 settings, the end user) to the producers of services / technologies involved in e-learning (programmers, administrators, teachers with particular innovative ideas). So mashup could be the ideal solution to Situational Application development, where we have a precise need of an integrated environment in which exploiting all the possibilities given by mashup philosophy.

We believe that, in the latest perspective, open and innovative development platforms like Eclipse could be the perfect candidate for this purpouse. In particular, mashup environments require modular and flexible architecture, allowing the users to manage every abstraction level of the mashup pyramid in a very simple way.

The field is still in its infancy, there is a lot of promising aspects but also dark points, especially from the end-user perspective: mashup could also be seen as a land of confusion, of unprecise construction, a fertile ground for chaos in learning objects and learning services production.

For this reason, further studies are requested especially from an experimental and technical point of view. For this purpose, a common project is also growing grouping together Napoli and Trento Universities with the aims to research newest mashup methodologies, technologies and best practices.

6. Bibliography.

[0] Anderson C. (2004), “*The Long Tail: Why the Future of Business Is Selling Less of More” (ISBN 1-4013-0237-8).*

[1] Maresca P. (2009) *La comunità eclipse italiana e la ricerca nel web3.0: ruolo, esperienze e primi risultati* Slide show for Mashup meeting at University of Salerno.

[2] Raimondo M. (2009) *Web 2.0 Modelli Sociali, Modelli di Business e Tecnologici* IBM Slide show for Mashup meeting at University of Salerno.

[3] IBM developerWorks Mashup section http://www.ibm.com/developerworks/spaces/mashups

[4] Duane Merril (2006) *Mashups: The new breed of Web app* http://www.ibm.com/developerworks/web/library/xmashups.html?S_TACT=105AGX01&S_CMP=LP

[5] IBM website www.ibm.com

[6]Jim Amsden, *Levels of Integration: five ways you can integrate with eclipse platform* , http://www.eclipse.org/articles/Article-Levels-Of-Integration/levels-of-integration.html

[7] Eclipe Rich Ajax Platform web site http://www.eclipse.org/rap/

[8] Eclipe Data Tooling Platform web site http://www.eclipse.org/datatools/

[9] Eclipe Business Intelligence and Reporting Tool web site http://www.eclipse.org/birt/

[10] Eclipe Rich Ajax Platform web site

http://www.eclipse.org/rap/

[11] Eclifox web site

http://www.alphaworks.ibm.com/tech/eclifox

[12] Elmeleegy, H. Ivan, A. Akkiraju, R. Goodwin, R., Mashup Advisor: A Recommendation Tool for Mashup Development, in: ICWS '08. IEEE International Conference on Web Services, 2008., Sept. 2008, Benjing, pp.337-344, ISBN: 978-0-7695-3310-0

[13] Marc Eisenstadt, "Does Elearning Have To Be So Awful? (Time to Mashup or Shutup)," icalt, pp.6-10, Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007), 2007

J-META: a language to describe software in Eclipse community

Pierpaolo Di Bitonto[(1)], Paolo Maresca[(2)], Teresa Roselli[(1)], Veronica Rossano[(1)], Lidia Stanganelli[(3)]
[(1)]Department of Computer Science – University of Bari
Via Orabona 4, 70125 Bari – Italy
[(2)]Dipartimento di Informatica e Sistemistica – Università di Napoli “Federico II”
Via Claudio 21, Napoli - Italy
[(3)]DIST - University of Genoa, Italy
Viale Causa 13, 16145, Genova, Italy
{dibitonto, roselli, rossano}@di.uniba.it, paolo.maresca@unina.it, lidia.stanganelli@poste.it

Abstract— **Information retrieval is one of the main activities in different domains such as e-commerce, e-learning or document management. Searching in large amount of data faces two main problems: the suitability of the results with respect to the user's request, and the quantity of the results obtained. One of the most popular solutions for this problem is to define more and more effective description languages. In order to allow the search engine to find the resource that best fit the user's needs, that can be very specific, detailed description are needed. Finding the right level of the description is the current challenge, of the researches in e-learning, e-commerce and document management domains. For instance, a teacher can search a Learning Object (LO) about a simulation of a chemical reaction, in order to enrich his/her courseware. Thus, the LO description should not contain only information about title, authors, time of fruition, and so on, but should contain more specific information such as the type of content, learning prerequisites and objectives, teaching strategy implemented, students addressed, and so forth. In an open source scenario, the problem is the same. In fact a developer needs to find a software component, which must be integrated in existing system architecture. In order to find the right component, technical issues should be described. It will be interesting if the open source web site supply an intelligent search engine which is able to select the software according with the developer's requirements. It will possible only if detailed descriptions of each resource are supplied. The paper proposes a description language, named J-META, to describe software resources (plug-in) in the Eclipse open community.**

Keywords-component; description language; Eclipse community; metadata

I. INTRODUCTION

Information retrieval is one of the main activities in different domains such as e-commerce, e-learning or document management. Searching in large amount of data faces two main problems: the suitability of the results with respect to the user's request, and the quantity of the results obtained. One of the most popular solutions for this problem is to define more and more effective description languages. Greater is the complexity of the resource which should be searched, greater is the complexity of the description. For example, in the WWW there are search engines which index the resources using their content (e.g. Google to describe web pages), and search engines that need extra details, such as title, authors, time, context, etc. (e.g. YouTube to describe multimedia resources).

In order to allow the search engine to find the resource that best fit the user's needs, that can be very specific, detailed description are needed. Finding the right level of the description is the current challenge of the researches in e-learning and document management domains. For instance, a teacher can search a Learning Object (LO) about a simulation of a chemical reaction, in order to enrich his/her courseware. Thus, the LO description should not contain only information about title, authors, time of fruition, and so on, but should contain more specific information such as the type of content, learning prerequisites and objectives, teaching strategy implemented, students addressed, and so forth.

In an open source scenario, the problem is the same. A developer needs to find a software component, which must be integrated in existing system architecture. In order to find the right component, technical issues (such as goals, functionalities, hardware and software requirements, relationship among other components, the kind of licence, etc.) should be described. Perhaps, these kind of information can be contained in the software documentation (if it exists), but the developer will spend a lot of time in reading and installing all software available in an open source web site. It will be interesting if the open source web site supply an intelligent search engine which is able to select the software according with the developer's requirements. It will possible only if detailed descriptions of each resource are supplied. Therefore, in order to allow the search engine to supply the best suitable resources for user's requests, it is necessary to work on large number of descriptors.

The paper proposes a description language, named J-META, to describe software resources (plug-in) in the Eclipse open community. *Eclipse is an open source community, whose projects are focused on building an open development platform comprised of extensible frameworks, tools and runtimes for building, deploying and managing software across the* *lifecycle. The Eclipse Foundation is a not-for-profit, member supported corporation that hosts the Eclipse projects and helps cultivate both an open source community and an ecosystem of complementary products and services* [1].

Since 2001, the Eclipse community has been growing quickly. Nowadays, the community counts: 180 members including Public Administration, small and big companies, universities and research centres; 66 software projects and the Eclipse platform is used as development platform in more than 1300 products. Eclipse is leader in Java development environment with about 2 thousand users over the world [2]. There are a lot of local communities and also an eclipse Italian community that manages about 10 projects.

The main idea of Eclipse community is to share ideas, knowledge and experiences. In this context, the J-META language presented herein, aims at defining a set of specifications which allows Eclipse community members to describe and find plug-in easily.

II. BACKGROUND

The state of art of metadata language has highlighted the lack of languages to describe software resources, in particular plug-in. So, to define J-META language it has been studied the description languages defined in other domain such as e-learning and librarian world in order to obtain useful guide lines and suggestions to the J-META language definition.

The studied languages (Dublin Core [3], IEEE Learning Object Metadata [4], Text Encoding Initiative [5], etc) have pointed out some advantages and disadvantages related to the description languages and their use in real contexts. For the sake of simplicity it will be described just two of these languages: Text Encoding Initiative (TEI) and IEEE Learning Object Metadata (LOM).

TEI language was born in librarian world to describe textual resources, and scholar text in particular; LOM language was defined in e-learning context to describe didactic resources.

The TEI standard has been developed many encoded data sets, ranging from the works of individual authors to massive collections of national, historical, and cultural literatures [5]. The TEI includes: (1) analysis and identification of categories and features for encoding textual data at many levels of detail; (2) specification of a set of general text structure definitions that is effective, flexible, and extensible; (3) specification of a method for in-file documentation of electronic text that is compatible with library, cataloging conventions and can be used to trace the history of the text and thus can assist in authenticating their provenance and the modifications; (4) specification of encoding convention for special kind text or text features (character sets, general linguistics, dictionaries, spoken texts, hypermedia).

The LOM standard is a set of IEEE specifications that serves to describe teaching resources or their component parts. It includes more than 80 descriptive elements subdivided into the following 9 categories:

- General: this includes all the general information that describes the resource as a whole. The descriptors in this group include: title, structure, aggregation level
- Lifecycle: this groups the descriptors of any subsequent versions of the LO and of its current state
- Meta-metadata: these include the information on the metadata themselves
- Technical: this indicates the technical requisites needed to run the LO and the technical characteristics of the LO itself, such as the format or size of the file
- Educational: this contains the pedagogical and educational features of the LO. This is the most important category and contains elements like: Interactivity type, Learning resource type, Semantic density, Typical learning time, that supply indications on how the resource can be used in the teaching program
- Rights: this indicates the intellectual property rights and any conditions of use of the LO, such as cost, as well as the information on copyright.
- Relation: this describes any relations (of the type: “is a part of”, “requires”, “refers to”) with other LOs.
- Annotation: this allows the insertion of comments about the use of the LO in teaching, including an identification of who wrote the annotation.
- Classification: this makes it possible to classify the LO according to a particular classification system.

The analysis of different description languages has point out two important suggestions for the definition of J-META. In particular, the TEI suggested the level of detail for the description of the plug-in and its each single component; the LOM suggested the data model.

III. J-META LANGUAGE

The state of art of metadata languages presented in the previous section arises two questions about the most important aspects for the description language definition: what kind of software should be described? Which grain size must be chosen to describe the software?

As concerning the kind of software that should be described (and then searched), it is chosen as minimum unit the plug-in because both functions (in the functional paradigm) and classes (in object oriented paradigm) have too much references with other functions or classes this makes difficult reuse the software in different contexts. On the contrary, the plug-in is independent from external code, in the worst case it depends from other plug-ins. As concerning the grain-size of plug-in descriptions it has been chosen a thin level describing functions or classes that compose the plug-in in order to describe with accuracy the software resources.

As figure 1 shown, J-META is composed by five main categories:

- General: the general information that describes the plug-in as a whole;
- Lifecycle: the features related to the history and the current state of the plug-in;
- Technical: the technical requirements and characteristics of the plug-in;
- Rights: the intellectual property rights and conditions of use for the plug-in;
- Code: the analysis and design diagrams and the code description of the plug-in.

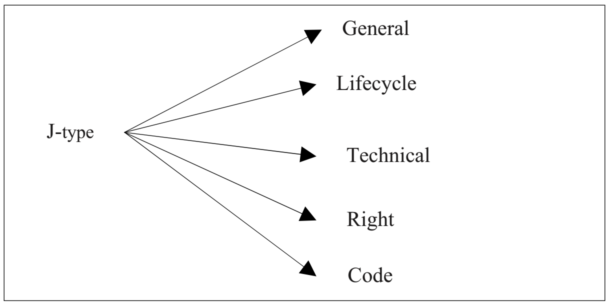


Figure 1. Main category of J-Meta language

Likewise, the LOM data model, J-Meta is a hierarchy of data elements, including aggregate data elements and simple data elements (leaf nodes of the hierarchy). The first level of hierarchy is subdivided in subcategories which describe the plug-in details. Each subcategories is composed by a number of data elements, each of those aiming at describe a particular plug-in issue. Only leaf nodes have individual values, for each single data element, J-Meta defines:

- Name: the name of the data element;
- Explanation: the definition of the data element;
- Size: the number of values allowed for the data element;
- Values: the set of allowed values for the data element;
- Type of data: indicates whether the values are String, Date, Duration, Vocabulary or String or Undefined.

For the sake of simplicity the data elements of each main category will be presented with a short description and examples in order to clarify their use in a real context.

A. *General category*

`General` category is composed by 8 data elements:

- `identifier:` a globally unique label that identifies the plug-in (for instance 0000AA, 00001, A3493);
- `title:` the name of the plug-in;
- `authors`: the name (or the names) of plug-in developer(s). It is an aggregate element that can contains one or more elements (`author`) to list all the authors involved;
- `description`: a description of the plug-in content;
- `keywords`: words (min 1 – max 10) that indicate technique or technology used in the development of plug-in (no controlled vocabulary is defined);
- `sector`: area of pertinence of plug-in, it describes the area of the service. The possible values are listed in a controlled vocabulary and are those suggested by the classification of plug-in central Eclipse site (Application Management, Application server, Build and Deploy, …)
- `annotations`: gives to the users the possibility of insert comments or recommendations about the use of plug-in. It is an aggregate data with the following children elements:
  - `name`: the name of the user that gives the comment;
  - `object`: object of the comment;
  - `comment`: body of the comment;
  - `rating`: user's global evaluation of the plug-in (expressed in numeric value from 1 up to 5)
- `download:` the number of downloads of the plug-in.

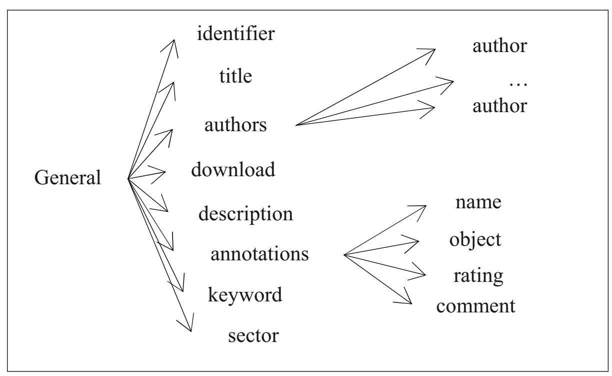


Figure 2. General category of J-Meta language

B. *Lifecycle category*

`Lifecycle` category is composed by 4 data elements

- `version`: alphanumeric string (max 10 characters) that indicates the version of the plug-in (for instance 3.2 or 3.2alpha)
- `state`: the state of progress of the plug-in, the possible values are (1) not-complete – the plug-in project has just been started and it needs to be completed, (2) draft – the plug-in is almost complete but it has not been tested, (3) complete – the plug-in

has been released after a test phase, (4) neglected – the plug-in is incomplete and the project closed.

- `releaseDate`: the release date of the plug-in; the format is mm/dd/yyyy;
- `contributes`: aggregate element used to track possible contributions or recommendations, like changes or improvements from other developers. Contributes element can have one or more children (`contribute`), all identified by the attribute contribute_id. Each contribute element has three children elements: (1) `contributor` that indicates the name of who modified the plug-in, (2) `data` that indicates when contribute was released; (3) `description` that describes the contribution.

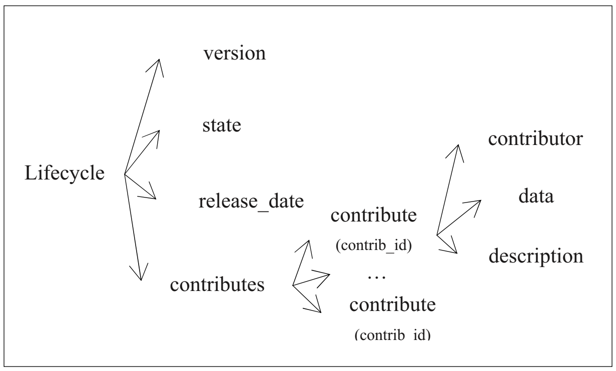


Figure 3. Lifecycle category of j-meta language

C. *Technical category*

The `technical` category is composed by 6 main elements:

- `size`: numeric value that expresses in bytes the size of the plug-in; this element will refer to the actual size of the plug-in, not to the compressed one
- `location`: a unique resource identifier in the Web, it may be a location (e.g., Universal Resource Locator), or a method that resolves to a location (e.g., Universal Resource Identifier)
- `eclipseRel`: the Eclipse release used to develop the plug-in (it uses a controlled vocabulary with the following values: Callisto, Europe, Ganimade, Galileo)
- `dependency`: the relationship between the plug-in and other plug-ins, it has a single child element (`resource`) that can occur one or more times, and describes the related plug-in. Children of resource element are: (1) `name` that identifies the name of the related plug-in, (i.e. j-viewer); (2) `reference` that indicates a unique resource identifier in the Web; (3) `description` that explains the relations among related plug-ins (i.e. the plug-in needs j-viewer to use file .jar);
- `requirements`: the technical capabilities (hardware and software) necessary for using the plug-in. Any constraints is expressed by `requirement` element (that can occur one or many times) that is a father node; its child are (1) `type` that is the type of the required technology to use the plug-in (i.e. hardware, software, network, etc); (2) `name` that is the name of the required technology to use the plug-in; (3) `minVersion` that is the lowest possible version of the required technology:
- `installationRemarks`: describes how to install the plug-in and the possible problems that could arise and their solutions;
- `documentationRepository`: pointers to link web where to find good documentation about plug-in. Web 2.0 tools are welcomed.

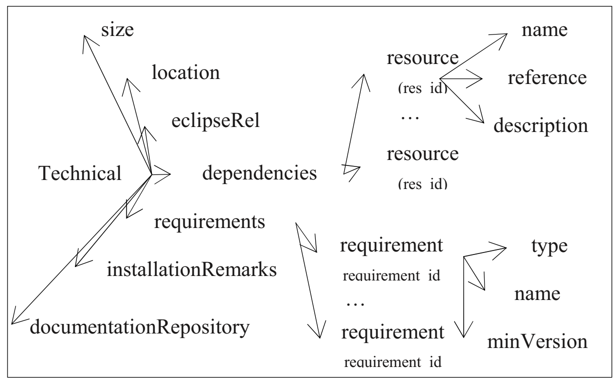


Figure 4. Technical category of j-meta language

D. *Rights category*

`The Rights` category is composed by three data elements:

- `cost:` indicates if the use of plug-in requires any payment (boolean values are only accepted);
- `licence:` the kind of software licence. The possible value are defined in a controlled vocabulary that contains the existing open software licences (GNU[6], ASL [7], BSD[8], …);
- `description:` comments about the plug-in conditions of use.

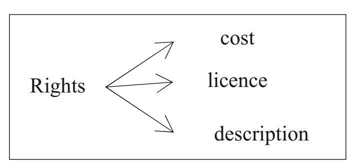


Figure 5. Rights category of J-Meta language

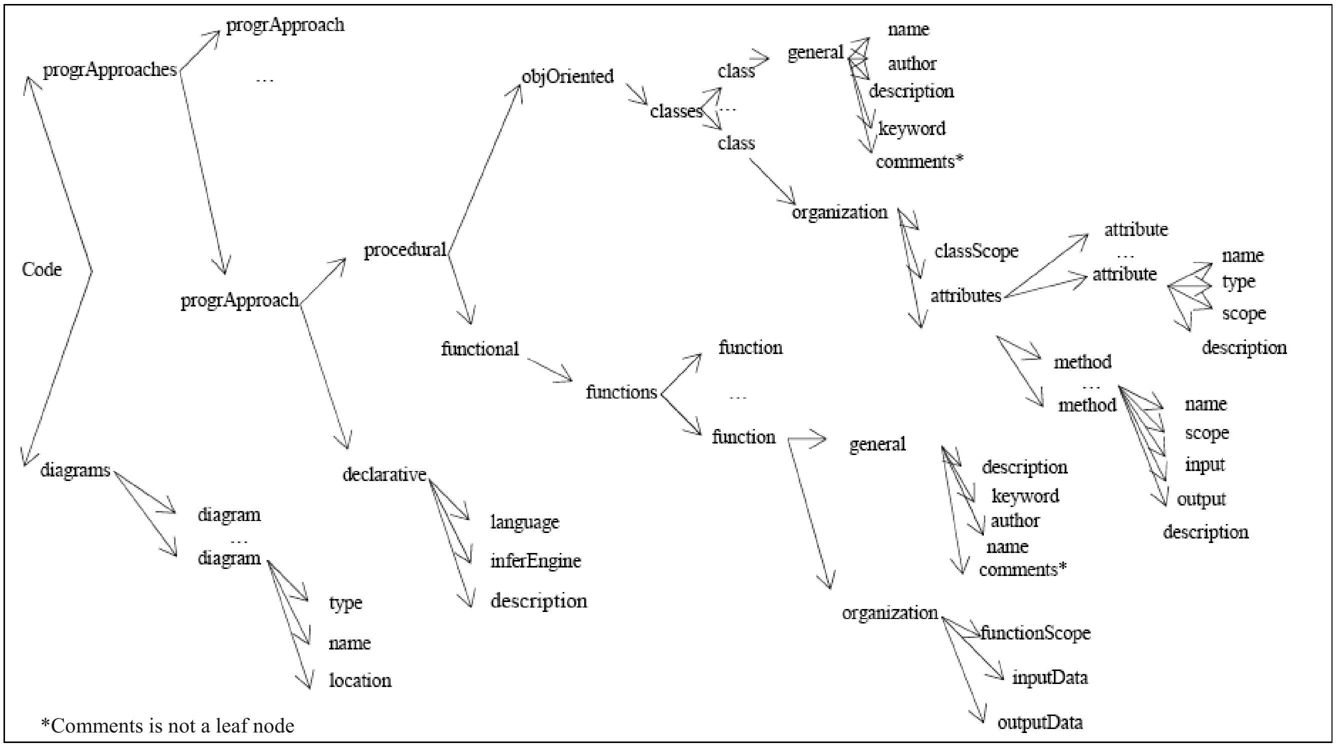


Figure 6. Code category of J-Meta language

E. Code category

`Code` category is composed by 40 data elements grouped in two main categories: `progrApproaches`, which describes the programming approach used (declarative, procedural, object oriented, functional, etc), and `diagrams`, which describes the code from a high abstraction level using UML diagrams that during analysis and design phases are defined. Diagrams category is defined to help programmers to better understand the source code. This is very important if a programmer needs to modify and/or extend an existent plug-in or needs to build a new plug-in that depends on an existing ones. These situations are very common in an open source community like Eclipse.From the source code point of view, the plug-in is a complex object that can use different programming paradigms, for instance Java code (procedural paradigm) and Prolog code (declarative paradigm), or C code (functional paradigm) and Java code (object oriented paradigm). This heterogeneity requires a flexible description structure adaptable to different contexts.

For this reason the data element `progrApproaches` has `progrApproach` element as child element that can occurs one or more times; each `progrApproach` element has `declarative` and `procedural` as children elements.

The declarative element describes the declarative approach used for the plug-in development. In particular, it uses three elements: (1) `language` that describes the programming language used (for instance Prolog, Clips [9], etc.); (2) `inferEngine` that specifies the inferential engine used for the plug-in development (for instance SWI Prolog [10] or SICStus prolog [11]); (3) `description` that describes the goals of the code and how it works.

The `procedural` element describes the procedural approach used for the plug-in development. In particular, it uses two different elements: `functional` and `objOriented` according to the different programming approaches. If functional paradigm is used, it is possible to describe, using the `functions` element, all the functions implemented.

In particular, using the function element it is possible to describe general information about the each single function and other information about its organization. General information (`general`) are: (1) `name`, the name of the function, (2) `author`, (3) `description`, (4) `keywords` and (5) `comments`. For each `comment`, it is possible to specify the user who makes the comment, the object, the body and the user's global evaluation (from 1 up to 5). The information about the organization of the function are described in:

`functionScope`, which indicates if the function is public or private; `inputData` and `outputData`, which describe the name and the type of input and output data of the function respectively.

If Object Oriented (OO) paradigm is used, J-meta can describe the `classes` used in the plug-in. It is possible to specify both general information such as (1) `name`, (2) `author`, (3) `description`, (4) `keywords`, (5) `comments`, and detailed information about the class and/or the plug-in organization. For what concerning the `class` organization it is possible to specify, using `classScope`, if the class is public or private, and its own attributes and methods. For each attribute, it is possible to describe the name of the attribute, its scope (public or private) and a textual explanation; for each method it is possible to specify the `name`, the `scope` (public or private), the name and type of input and output data and a short explanation.

IV. CONCLUSIONS AND FUTURE WORKS

The paper proposes a description language, named J-META, to describe software resources (plug-in) in the Eclipse open community with the aim at improving the accuracy of the search process. The main problems of the searching activities in Eclipse Community are the large amount of plug-ins and the complexity of the resources.

The precision of searching is strictly connected to the description of the resource that will be searched. Better detailed is the description of the resource, higher is the precision of the search. On the basis of these premises the language J-META has been defined. J-Meta allows to describe the software plug-in from different points of view such as goals, functionalities, hardware and software requirements, relationship among other components, kind of licence, etc. From the structural point of view, J-Meta is a hierarchy of data elements, including aggregate data elements and simple data elements (leaf nodes of the hierarchy). The first level of hierarchy is subdivided in subcategories which describe the plug-in details. Each subcategories is composed by a number of data elements, each of those aiming at describe a particular plug-in issue.

In the Eclipse Community does not exist any description language for the plug-in, so the developers are forced to read the different plug-in documentation in order to find the resource that best fit to their needs. The introduction of J-Meta will improve the search engine performances and the users' satisfaction.

The next step will be the validation of the J-META within the Eclipse Italian Community, and then in the worldwide Eclipse Community.

REFERENCES

[1] http://www.eclipse.org

[2] Paolo Maresca. Project and goals for the Eclipse Italian Community. (2008) International Workshop on Distance Education Technologies (DET'2008), part of the 12th International Conference on Distributed Multimedia Systems, Boston, USA, 4-6 September, 2008.

[3] Dublin Core Metadata Initiative (DCMI) http://dublincore.org/

[4] IEEE Learning Object Metadata (LOM) http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf

[5] Ide, N. and Sperberg-McQueen, C. The TEI: History, goals, and future. Computers and the Humanities 29, 1 (1995), 5–15.

[6] GNU General Public License http://www.gnu.org/copyleft/gpl.html

[7] Apache License, Version 2.0 http://www.apache.org/licenses/LICENSE-2.0.html

[8] Open Source Initiative OSI - The BSD License:Licensing http://www.opensource.org/licenses/bsd-license.php

[9] Clips http://clipsrules.sourceforge.net/

[10] SWI Prolog http://www.swi-prolog.org/

[11] Sicstus Prolog http://www.sics.se/isl/sicstuswww/site/index.html

Providing Instructional Guidance with IMS-LD in COALA, an ITS for Computer Programming Learning

Francisco Jurado, Miguel A. Redondo, Manuel Ortega
University of Castilla-La Mancha
Computer Science and Engineering Faculty
Paseo de la Universidad 4
13071 Ciudad Real, Spain
+34 295300
{Francisco.Jurado, Miguel.Redondo, Manuel.Ortega}@uclm.es

Abstract—**Programming is an important competence for the students of Computer Science. These students must acquire knowledge and abilities for solving problems and it is widely accepted that the best way is learning by doing. On the other hand, computer programming is a good research field where students should be assisted by an Intelligent Tutoring System (ITS) that guides them in their learning process. In this paper, we will present how we have provided guidance in COALA, our ITS for programming learning, merging Fuzzy Logic and IMS Learning Design.**

Keywords-Problem Based Learning, Intelligent tutoring, Adaptive environments, Instructional planning

I. INTRODUCTION

Obtaining Computer Programming competence implies that students of Computer Science should acquire and develop several abilities and aptitudes. It looks to be widely accepted that the best way to acquire that competence is learning by doing [1]. Therefore, students must solve programming problems presented strategically by the teacher. These students use computers in order to develop the learning activities which the teacher has specified. Thus, we think this makes it an ideal environment for Computer Assisted Learning (CAL), where students are assisted by an Intelligent Tutoring System (ITS) that guides them in their learning process, helping them to improve and to acquire the abilities and aptitudes they should acquire and develop, leaving behind the slow trial and error process.

An ITS allows adapting the learning process to each student. For this, the ITS is based on determining what the student cognitive model is, so it can determine the next learning activity for each specific student. ITS are usually used together with Adaptive Hypermedia Systems (AHS) for providing "intelligent" navigation through the educative material and learning activities. These systems that merge ITS and AHS are the so-called Adaptive Intelligent Educational Systems (AIES). Several examples of systems that integrate AHS and ITS for programming learning can be found in EML-ART [4], Interbook 3, KBS-Hyperbook [13] or AHA! [5].

On the other hand, the growth of Web-Based Education (WBE) environments has made groups such as IEEE LTSC, IMS Global Learning Consortium, or ADL, work to provide a set of standards allowing reusability and interoperability into the eLearning industry, setting the standard base where engineers and developers must work to achieve eLearning systems integration.

One of those specifications is IMS Learning Design (IMS-LD) [8] proposed with the aim of centering on cognitive characteristics and in the learning process, allowing isolating the learning design process from the learning object development.

In this paper, we will show how we have extended our distributed environment for Programming Learning called COALA (Computer Assisted Algorithm Learning), integrating an IMS-LD engine into it. This approach merges, on the one hand, AHS techniques used in the system mentioned previously by using eLearning standards specification, and on the other hand, Artificial Intelligent Techniques for enabling ITS to lead students to achieve the abilities and aptitudes they need for their future work.

The paper is structured as follows: firstly, an overview about what we must take into account for providing instructional adaption (section 2); secondly, an explanation of our instructional model (section 3); then, the assessment or evaluation service the system uses will be shown (section 4); next, we will go deeply into some implementations issues and how the system works (section 5); finally, some concluding remarks will be made (section 6).

II. OVERVIEW

Our aim is to provide an approximation that allows the creation of ITS considering student cognitive model and the instructional strategy needed to teach the lesson. In this sense, it is necessary to use techniques from AHS, as summarized in Brusilovsky's taxonomy [2]:

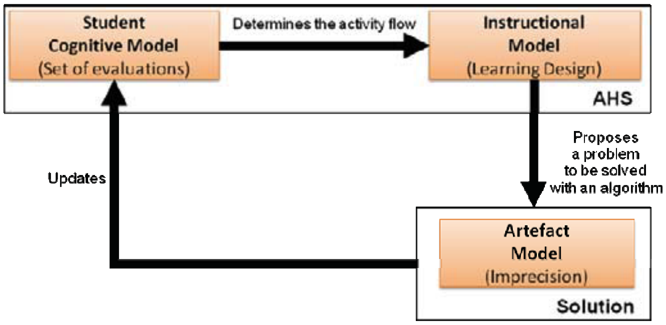


Figure 1. Models in our Approach.

- A user model based on concepts: this consists of a set of concepts with attributes such as the degree of knowledge. Then, for instance, in AHA! [5], visiting a webpage implies incrementing the knowledge attribute for the concept dealt with in that webpage. Also, updating the knowledge attribute can be propagated to other concepts.
- Adaptive link hiding: this means that a set of Boolean expressions can be defined based on values from the user model. With this, the showing and hiding of a link can be evaluated.
- Conditional inclusion of fragments: this introduces a set of conditional blocks that allows the appearance or not of text fragments.

In instructional design, we must provide learning activities sequencing. Thus, Brusilovsky's taxonomy will be adopted in our proposal, defining some models which are: the student cognitive model, the instructional model and the artifact model [11].

On the top left-hand corner of the figure 1, we can see the student cognitive model. It consists of a set of evaluations for each task the student has to solve and represents the cognitive stage for the student at every moment. This matches the user model, based on the concept taken from Brusilovsky's taxonomy. In other words, it specifies what parts of the domain the student knows and to what degree.

On the top right-hand corner, the figure shows the instructional model. This model allows specifying the instructional strategy or learning design to be applied. In other words, the instructional model represents the learning activity flow. It will be adapted depending on the evaluations stored in the student cognitive model. This matches the adaptive link hiding and conditional inclusion fragments from Brusilovsky's taxonomy. Thus, learning activities substitute links and fragments, for example reading a text, designing quizzes, multimedia simulation, chats, etc.

Among the learning activities, a problem to be solved with an algorithm can appear. These algorithms should be analyzed, assessed and evaluated. Then, to support this, a model that manages the solution must be considered. This will be the artifact model shown in the bottom right-hand corner of figure 1. This model allows supporting processing and analyzing artifacts (algorithms) developed as a solution for a proposed problem. This must provide a mechanism for managing the imprecision and vagueness with which both teacher and student specify the solution.

This artifact model, which analyzes the solution, interacts with the student cognitive model for updating it, reflecting the evidence of knowledge that has given shape to the solution developed by the student. In this way, in accordance with the student's work, the instructional adaptation can be achieved, deciding the next learning activity to be proposed. Thus, the learning activities are shown as a consequence of how the student solves problems.

So, for implementations proposal, we use IMS-LD in our instructional model, and Fuzzy Logic [16] in our artifact model for the evaluation process [9] [10]. In the following sections, we will show in detail how we have implemented these models into our system.

III. IMS-LD FOR THE INSTRUCTIONAL MODEL

As we have previously stated, we need learning activity sequencing to set our instructional model. Since our aim is to develop an ITS that allows applying instructional strategies according to the subject to be learned/taught, we propose the use of IMS-LD [8] to specify the method that must be used in the teaching/learning process [14], that is, to specify the instructional adaptation model.

IMS-LD can be used for developing adaptive learning (AL) [15]. This is because an IMS-LD specification can be enriched with a collection of variables from the student profile. These variables allow specifying conditions to determine if a learning activity or a branch of the learning flow (a set of learning activities) is shown to or hidden from a specific student. It can be done in the following way: each method has a section for defining conditions that points out how it must adapt itself to specific situations through rules like those shown in figure 2. In that example code, if the student knowledge about a concept is less than 5, then the activity A1 is hidden and the activity A2 is shown; in the opposite case, activity A1 is shown and activity A2 is hidden.

In our system, the variables used for defining the adaptation rules in the condition section of an instructional method are obtained from the student model. In our case of study (programming learning), the evidence must obtain its value from the artifact (algorithm) developed by the student. In the next section, we explain how to evaluate the algorithms that students design as a result of the learning activities and how their cognitive model is updated.

```
IF student knowledge less-than 5
  THEN hide activity A1 and show activity A2.
  ELSE show activity A1 and hide activity A2.
```

Figure 2. Example of the Rule for Adapting an Instructional Design.

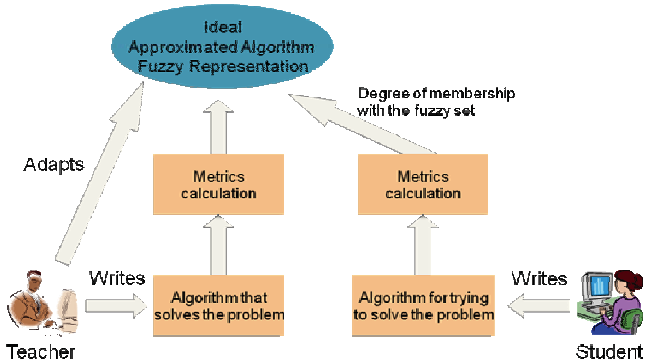


Figure 3. Evaluating the Student Algorithm.

IV. THE ARTIFACT MODEL FOR EVALUATING THE STUDENT SOLUTION

Up to this point, we have explained how a generic ITS can be implemented, taking into account the student model and the instructional model which must be followed. In our case, we want to apply this ITS to a programming learning environment. Thus, the learning activities that the IMS-LD player will show to students can be programming problems. So, a model that evaluates the algorithm delivered as a solution is necessary.

Our proposal is explained in [9] [10] and briefly shown in figure 3. In this figure, the teacher writes an implementation for the ideal approximate algorithm that solves a problem (on the bottom left of the figure). Next, several software metrics that shape its functionality will be calculated. In this way, we obtain an instance of the ideal approximated algorithm. After that, the fuzzy set for each metric will be established in the following way: initially, each fuzzy set will be a default trapezoidal function around the metric value from the approximate algorithm; he teacher can adapt each fuzzy set for indicating whether an algorithm is correct or not. From this, we obtain a collection of fuzzy sets that characterize the algorithm. Thus, we get a fuzzy representation of that ideal approximated algorithm, that is, we obtain an ideal approximated algorithm fuzzy representation that solves a concrete problem (at the top of figure 3).

Algorithms that students have written (on the right of the figure) will be correct if they are instances of that ideal algorithm fuzzy representation. Knowing the degree of membership for each software metric, obtained from the algorithm written by students in the correspondent fuzzy set for the ideal approximated algorithm fuzzy representation, will give us an idea of the quality of the algorithm that students have developed.

With this method, we have an artifact model that manages imprecision and vagueness; furthermore, it is based on solid engineering practice (software engineering).

Thus, the system will have the evaluation of the algorithm developed by the student as feedback. This can be used by the teacher for re-writing or adapting both the learning design and the approximated ideal fuzzy representation of the algorithm in order to improve the system.

V. IMPLEMENTATION ISSUES

For the implementation, we have taken COALA (Computer Assisted Environment for Learning Algorithms) as a starting point [12]. COALA has been developed as a customized Eclipse application by means of plug-ins. It is an Integrated Development Environment (IDE) that is not so different from the one that students will find in their future work. That is, it doesn't use virtual environments or simulation tools, but employs a real-world IDE. COALA allows the distribution of programming tasks or assignments to students, the asynchronous downloading of such assignments, local elaboration, uploading, annotation and feedback to teachers and students.

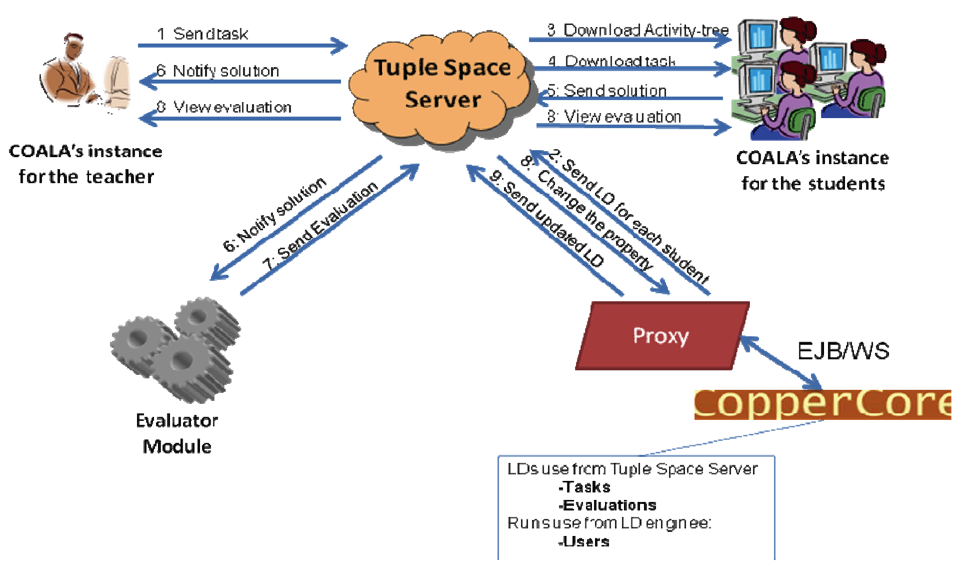


Figure 4. IMS-LD as a Service.

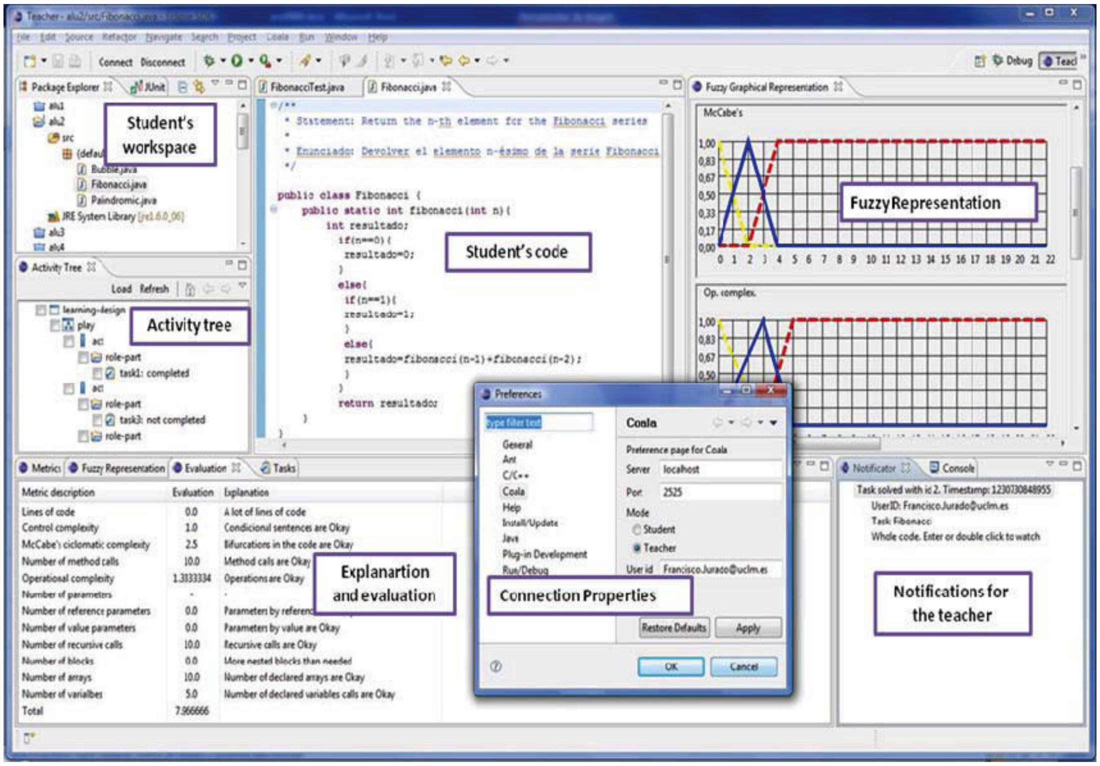


Figure 5. Customized Eclipse Framework.

As a communication engine, COALA implements Blackboard architecture by using a Tuple Space (TS) server. A TS server is basically a shared memory in which clients can store and retrieve information formatted as tuples [6]. The COALA plug-in for the Eclipse environment allows communication by means of a TS implementation called SQLSpaces, developed at the University of Duisburg-Essen [7].

To provide the corresponding guidance, we have chosen the main LD engine which can be found nowadays: CopperCore. As communication middleware, CopperCore uses Enterprise Java Beans (EJB) or Web Services (WS). So we have implemented a proxy that translates the necessary API and allows communication with our Tuple Space server.

Following our explanation, figure 4 shows the different steps and messages (tuples) between the teacher, the students, the evaluator module and the LD engine by means of the TS server. So, as we can see in figure 4, in the beginning, the teacher specifies an assignment using his/her COALA environment (figure 5). After this, the teacher uploads this assignment by sending a tuple with the form <task id, task description, test cases, fuzzy representation> (step 1) using the "Send Task to TS" action in the plug-in. At that moment, the task is available to all the students in the classroom.

The tasks the teacher had uploaded must be those specified in the corresponding IMS-LD which has been previously loaded in CopperCore. On the other hand, the LD engine uses the proxy to talk to the TS server making use of the tasks stored on it. Then, CopperCore can send the TS server a tuple with the form <user id, run id, activity tree> which contains the corresponding activity tree extracted from the IMS-LD specification for each student (step 2). At this time, all the tasks and the corresponding activity trees for each student are available.

Therefore, the students can, first of all, download their activity tree specification using the "Learning design" view in their COALA environment (figure 5, on the left) reading the <user id, run id, activity tree> tuple (step 3). Secondly, following this activity tree, the students are able to download the corresponding assignment onto their workspace reading the tuple <task id, task description, test cases, fuzzy representation> previously uploaded by the teacher (step 4), using their "Download Task from TS" action menu in the plug-in. Then, each student can work out the task by writing the code, compiling, etc.

Once the students have finished the assignment, they can send their results to the server from where they can be downloaded and reviewed by the teacher. Students upload the solution to the server sending a tuple with the following content: <use id; task id; solution code> (step 5). The teacher will be notified about the task sent, and can check the code written by the student on his/her computer by reading all the tuples with the form <use id; task id; solution code> (step 6) from the server. Now the teacher can see the task in his/her "Notification view" (figure 5, on the bottom right).

As previously mentioned, we have implemented an evaluator module that reads the tuples the students have sent, that is, the same ones the teacher reads (step 6), and processes the code for obtaining a set of metrics and an evaluation explanation (as presented in section 3). These calculated metrics are sent to the Tuple Space server with the form <task id; user id; metric1; metric2; ... metricN>. Also, an explanation associated with each metric is sent in a tuple with the following format: <task id; user id; explain metric1; explain metric2; ... explain metricN> (step 7). Then, both the teacher and the students can read the software metrics and the corresponding explanations from the server and analyze them (step 8). So, during their programming, students can use the tests created by the teacher and ask the system for an automatic evaluation to check their solution (figure 5, on the bottom left).

At the same time, (step 8) the proxy reads the evaluation for the task and informs CopperCore that a property has changed for a user. Then, CopperCore processes this change and updates the activity tree for the concrete user. The update in the activity tree tuple fires a notification to the student COALA instance. This notification informs COALA that the new activity tree is available and it will be downloaded as in step 4. So, the student can follow his/her activity tree and download the next task.

VI. CONCLUDING REMARKS AND FUTURE WORK

Throughout this paper, we have shown how we have created an ITS by merging techniques from the AHS and AI techniques. The paper starts by analyzing the necessary techniques from the AHS. We have adopted these AHS techniques by means of a set of models which are: the student cognitive model for determining which parts of the domain the student knows, the instructional model for adapting the learning activities flow depending on the knowledge the student has, and the artifact model for evaluating the students' solutions to assignments. This last model will update the student cognitive model as a consequence of the learning activity flow.

So, starting from our distributed environment, called COALA (Computer Assisted Environment for Learning Algorithms), which enables the distribution, monitoring assessment, and evaluation of assignments, we have shown how we have added, without difficulty, CopperCore as an IMS-LD engine. This was possible by integrating a new component in its blackboard distributed architecture, proving the extension capabilities of our architecture.

Thus, we have an ITS that allows adapting the learning process to each student, taking into account the results of the delivered assignment. So, as future work we intend to test the system in scenarios where an adaptation is needed, and then check if the system provides the correct one.

VII. ACKNOWLEDGEMENTS

This research work has been partially supported by the Junta de Comunidades of Castilla-La Mancha, Spain, through the projects AULA-T (PBI08-0069) and M-GUIDE (TC20080552) and Ministerio de Educación y Ciencia, Spain, through the project E-Club-II (TIN-2005-08945-C06-04).

REFERENCES

[1] Bonar, J. & Soloway, E., "Uncovering Principles of Novice Programming", in POPL '83: Proceedings of the 10th ACM SIGACT-SIGPLAN Symposium on Principles of Programming Languages, ACM, 1983, pp. 10-13.

[2] Brusilovsy, P., "User Modeling and User-Adapted Interaction", in Adaptative Hypermedia Vol. 11, nr. 1-2, Kluwer Academic Publisher, 2001, pp. 87-110

[3] Brusilovsy, P., Eklund, J., and Schwarz, E., "Web-Based Education for All: A Tool for Developing Adaptative Courseware", in Proceedings of Seventh Internacional World Wide Web Conference, 1998, pp. 291-300

[4] Brusilovsky, P., Schwarz, E.W. and Weber, G., "ELM-ART: An Intelligent Tutoring System on World Wide Web", in Intelligent Tutoring Systems, 1996, pp. 261-269

[5] De Bra, P., Ad Aerts, Bart Berden, Barend de Lange, Brendan Rousseau, "AHA! The Adaptative Hypermedia Architecture", in Proceeding of HT'03, 2003

[6] Gelernter, D., "Generative Communication in Linda", ACM Transactions on Programming Languages and Systems, 7(1), 1985, pp 80-112

[7] Giemza, A., Weinbrenner, S., Engler, J., Hoppe, H.U., "Tuple Spaces as Flexible Integration Platform for Distributed Learning Environments", in Proceedings of ICCE 2007, Hiroshima (Japan), November 2007. pp. 313-320

[8] IMS-LD, "IMS Learning Design. Information Model, Best Practice and Implementation Guide, XML Binding, Schemas. Version 1.0 Final Specification", Technical report, IMS Global Learning Consortium Inc, Online, http://www.imsglobal.org/learningdesign/index.cfm., 2003

[9] Jurado, F.; Redondo, M. A. & Ortega, M., "Fuzzy Algorithm Representation for its Application in Intelligent Tutoring Systems for the Learning of Programming", in Rogério PC do Nascimento; Amine Gerqia; Patricio Serendero & Eduardo Carrillo (ed.), EuroAmerican Conference On Telematics and Information Systems, EATIS'07 ACM-DL Proceeding, Association for Computing Machinery, Inc (ACM), Faro, Portugal, 2007

[10] Jurado, F.; Redondo, M. A. & Ortega, M., "Applying Approximate Reasoning Techniques for the Assessment of Algorithms in Intelligent Tutoring Systems for Learning Programming" (in Spanish), in Isabel Fernandez de Castro (ed.), VII Simposio Nacional de Tecnologías de la Información y las Comunicaciones en la Educación (Sintice'07), Thomson, Zaragoza, Spain, 2007, pp. 145-153

[11] Jurado, F.; Redondo, M. A. & Ortega, M., "An Architecture to Support Programming Algorithm Learning by Problem Solving", in Emilio Corchado; Juan M.Corchado & Ajith Abraham, (ed.), Innovations in Hybrid Intelligent Systems, Proceedings of Hybrid Artificial Intelligent Systems (HAIS07), Springer Berlin Heidelberg New York, Salamanca, Spain, 2007, pp. 470-477

[12] Jurado, F.; Molina, A. I.; Redondo, M. A.; Ortega, M.; Giezma, A.; Bollen, L. & Hoppe, H. U., "COALA: Computer Assisted Environment for Learning Algorithms", in J. Ángel Velázquez-Iturbide; Francisco García & Ana B. Gil, (ed.), X Simposio Internacional de Informática Educativa (SIIE'2008), Ediciones Universidad Salamanca, Salamanca (España), 2008

[13] Nejdl, W. and Wolper, M., "KBS Hyperbook—A Data Driven Information System on the Web", in WWW8 Conference, Toronto, 1999

[14] Oser, F.K. and Baeriswyl, F.J., "Choreographies of Teaching: Bridging Instruction to Learning", in Richardson, V. (ed.), Handbook of Research on Teaching. 4th Edition. McMillan - New York, 2000, p. 1031-1065

[15] Towel, B. and Halm, M., Learning Design: A Handbook on Modelling and Delivering Networked Education and Training, Springer-Verlag, chapter 12 - Designing Adaptive Learning Environments with Learning Design, 2005, pp. 215-226

[16] Zadeh, L. , "Fuzzy Sets", in Information and Control, Vol. 8, 1965, pp. 338-358

Deriving adaptive fuzzy learner models for Learning-Object recommendation

G. Castellano, C. Castiello, D. Dell'Agnello, C. Mencar, M.A. Torsello
Computer Science Department
University of Bari
Via Orabona, 4 - 70126 Bari, Italy
castellano, castiello, mencar, danilodellagnello, torsello@di.uniba.it

Abstract

Adaptive e-learning systems are growing in popularity in recent years. These systems can offer personalized learning experiences to learners, by supplying each learner with learning contents that meet his/her specific interests and needs. The efficacy of such systems is strictly related to the possibility of automatically deriving models encoding the preferences of each learner, analyzing their navigational behavior during their interactions with the system. Since learner preferences may change over time, there is the need to define mechanisms of dynamic adaptation of the learner models so as to capture the changing learner interests. Moreover, learner preferences are characterized by imprecision and gradedness. Fuzzy Set Theory provides useful tools to deal with these characteristics. In this paper a novel strategy is presented to derive and update learner models by encoding preferences of each individual learner in terms of fuzzy sets. Based on this strategy, adaptation is continuously performed, but in earlier stages it is more sensitive to updates (plastic phase) while in later stages it is less sensitive (stable phase) to allow Learning-Object suggestion. Simulation results are reported to show the effectiveness of the proposed approach.

1 Introduction

In the age of knowledge, e-learning represents the most important and revolutionary way to provide educational services at any time and place. Today, in each kind of learning environment, the learner covers a key role: he became the main protagonist of his learning pathway opening a new challenge for current systems that have to necessarily adapt their services to suit the variety of learner needs [1]. This trend has led to the development of user-centred e-learning systems where the main aim is to maximize the effectiveness of learning by supplying an individual learner with personalized learning material [5, 6]. In this kind of systems, also known as adaptive e-learning systems, personalization plays a central role devoted to tailor learning contents according to the specific interests of learners in order to provide highly personalized learning sessions [4]. To achieve this aim, individuality of each learner has to be taken into account so as to derive a learner model that encodes his/her characteristics and preferences. The derived learner model can be successively exploited to select, among the variety of available Learning-Objects (LOs), those that match the interests of the individual learner. Therefore, in order to develop an adaptive e-learning system, two main activities should be carried out: (i) the automatic derivation of learner models starting from the information characterizing the preferences of each learner and (ii) the recommendation of LOs on the basis of the learner model previously derived.

Typically, learner models are derived through the analysis of the navigational behavior that each learner exhibits during his/her interactions with the system. Obviously, the interests and needs of learners may evolve during the learning process. This is an important aspect that has to be considered in order to derive learner models that may be adapted over time so as to capture the changing needs of each learner [3].

In addition, learner preferences are heavily permeated by imprecision and gradedness. In fact, learner interests have a granular nature for which they cannot be referred to specific LOs but, rather, they cover a range of LOs somehow similar (e.g. typically a learner may prefer one or more LOs about similar or related topics). Moreover, learner characteristics apply to learning resources with graduality, that is, a characteristic applies to a LO on the basis of a compatibility degree. In other words, there is a compatibility degree between learner preferences and LOs which may vary gradually. As an example, the interest of a learner for "Computer Science" may apply to a LO about "Web" and to a LO concerning "Computer Architecture" with different compatibility degrees.

A mathematical framework suitable to represent and handle such imprecision and gradedness is Fuzzy Set Theory (FST) [8, 10], based on the idea of fuzzy sets, that are basic elements suitable for representing imprecise and gradual concepts. FST provides fuzzy operators that can be used to combine, aggregate and infer knowledge from fuzzy sets.

In this work, we propose a strategy that derives learner models representing learner preferences of each learner in terms of fuzzy sets. The strategy is able to dynamically adapt models to the changing learner preferences so as to recommend similar LOs at the next accesses of a learner. The adaptation of learner models is performed continuously via a process that comprises two phases: a plastic phase, during which the adaptation process is more sensitive to updates, and a stable phase, in which adaptation is less sensitive so as to allow LO recommendation. The two-phase adaptation process guarantees the convergence to a learner model that can be used to suggest new LOs that are compatible with the specific learner preferences.

The paper is organized as follows. In section 2 the approach proposed for modeling learners is briefly described, along with the basic mechanism used to associate LOs to learners, according to their preferences. In section 3 the strategy for adaptation of learner models formalized. Section 4 shows some simulation results to prove the effectiveness of the proposed approach. Finally, section 5 closes the paper by drawing some conclusions.

2 The proposed approach

The main idea underlying our approach is to describe a learner model using the same representation used to describe LOs. This provides a straightforward mechanism to recommend LOs to users on the basis of a compatibility degree. The common representation shared between learner models and LOs is based on metadata describing specific attributes. Unlike conventional metadata specifications, that assume attribute values to be precise (crisp), we allow attribute values to be vague (fuzzy) by using a representation based on fuzzy sets.

The theory of fuzzy sets [8] basically modifies the membership concept: a set is characterized by a membership function that assigns to each object a grade of membership ranging in the interval [0,1]. In this way, fuzzy sets allow a partial membership of their elements and they are appropriate to describe vague and imprecise concepts. The use of fuzzy sets together with particular mathematical operators defined on fuzzy sets provides a suitable framework for handling imprecise information. Since LO's (and learner's) attributes may be vague and imprecise, the employment of fuzzy sets to define their values can be of valuable help, leading to a description based on the so-called fuzzy metadata. Fuzzy metadata provide a general description of attributes related to a LO, characterized by both precise and vague properties. In particular, using fuzzy metadata, we can formalize the following kinds of attributes:

- attributes with crisp values, such as the Dimension (expressed in KB) of a LO;
- attributes with collective values, such as the Topic of a LO, which can assume categorical values (e.g. "Computer Science", "Economy", "Business", ...);
- attributes with imprecise values, such as the Fruition Time required by a LO, which can be expressed by a vague concepts such as LOW, MEDIUM or HIGH.

One key feature of the proposed model is the possibility to easily formalize imprecise properties, thus favoring a mechanism of information processing that is in agreement with human reasoning schemes [9].

In the following subsections, the description of both LOs and learner models is detailed.

2.1 Description of Learning-Objects

Each LO is defined by a collection of fuzzy metadata, i.e. a set of couples $<attribute, fvalue>$ where $attribute$ is a string denoting the name of an attribute and $fvalue$ is a fuzzy set defined on the domain of the attribute. An example of fuzzy metadata is:

$$\langle Complexity, \{Low/1.0, Medium/0.8, High/0.2\}\rangle$$

which means that the attribute *Complexity* is defined by a fuzzy set comprising three fuzzy values: Low that characterizes (belongs to) the attribute with membership degree equal to 1.0; $Medium$ with membership degree equal to 0.8 and $High$ with membership degree equal to 0.2.

More formally, denoted by A the set of all possible attributes, we define a fuzzy metadata as a couple $\langle a, \mu\rangle$ being $a \in A$ an attribute and $\mu : Dom\,(a) \longrightarrow [0,1]$ a fuzzy set defined on $Dom(a)$ (the set of all possible values of attribute a). Then, a learning resource LO is described by a set of fuzzy metadata, i.e.

$$LO = \{\langle a, \mu\rangle\, | a \in A\} \qquad (1)$$

with the constraint that each attribute occurs at most once in the description:

$$\forall \langle a', \mu'\rangle\,, \langle a'', \mu''\rangle \in LO : a' = a'' \rightarrow \mu' = \mu'' \qquad (2)$$

A very simple example of LO description is reported in figure 1. It can be seen how fuzzy metadata extend classical metadata since they can describe precise as well as imprecise properties characterizing the attributes of a LO. The attribute "Title", for example, has a crisp nature, hence it is represented as a singleton fuzzy set "Java basis course"

$\langle$ Title, { Java basis course / 1.0 }$\rangle$
$\langle$ General topics, { Computer Science /1.0 }$\rangle$
$\langle$ Specific topics, { Programming / 0.8, Operating Systems / 0.2 }$\rangle$
$\langle$ Complexity, { LOW / 1.0, MEDIUM / 0.8, HIGH / 0.4 }$\rangle$
$\langle$ Fruition time, { Trapezoidal(15,30,60,90) }$\rangle$

Figure 1. An example of LO description

$\langle$ Topics, { Fuzzy Set Theory / 1.0, Neural Networks / 0.8 }$\rangle$
$\langle$ Genre, { theoretical / 1.0, applicative / 0.1, survey / 1.0 }$\rangle$

$\langle$ Topics, { C++ / 0.2, Java / 0.8, Smalltalk / 0.3 }$\rangle$
$\langle$ Target, { researcher / 0.1, undergraduate / 1.0, graduate / 0.5 }$\rangle$

Figure 2. An example of learner model with two components

with full membership degree. The attribute "Specific topics" is characterized by collective values, hence it is described by a fuzzy set enumerating two values: Programming, with membership degree 0.8, and Operating systems with membership degree 0.2. This means that this LO deals mainly on Programming, and, to a lesser extent, on Operating Systems. The attribute "Complexity" has a granular nature, hence it can be defined by enumerating three values (LOW, MEDIUM and HIGH) with different membership degrees. Finally, the attribute "Fruition time" has an imprecise and continue nature, hence it is described by a fuzzy set characterized by a trapezoidal membership function defined on the domain of time (expressed in minutes).

2.2 Description of learner models

Learner models are used to represent the preferences of each individual learner accessing the e-learning system. Precisely, a learner model reflects the preferences the learner has for one or more attributes of the accessed LOs. We define a learner model as a collection of components, where each component represents an elementary preference that is characterized in terms of fuzzy sets, likewise the fuzzy metadata specification used for LO description. This homogeneity enables a very direct matching between model components and LOs, so as to derive a compatibility degree useful for LO recommendation.

Formally, a learner model is defined as:

$$P = \{p_1, p_2, \ldots\} \tag{3}$$

where each component p_i is represented using a LO description, i.e.

$$p_i = \{\langle a, \mu\rangle \,| a \in A\} \tag{4}$$

A learner model is initially empty (i.e. it has zero components), then it incrementally grows by adding a component or updating the existing components each time the learner accesses a new LO. This dynamic adaptation of learner models is described in section 3.

In fig. 2, an example of learner model with two components is reported. We may interpret this model as a learner with two different types of interests. The first component indicates that the learner is interested mainly on Fuzzy set theory and, to a lesser extent, on Neural networks. The second component indicates that the same learner is mainly interested in Java and, to a minor extent, in Smalltalk and C++. It also indicates that preferred LOs should be mainly targeted to undergraduate students, while LOs targeted to researchers and graduate students are not of main interest for this learner.

2.3 Matching mechanism

Given a Learning-Object description LO defined as in (1) and a learner model P defined as in (3), we define a matching mechanism to compute a compatibility degree between LO and P that is as high as the learning resource is deemed compatible with learner's interests and preferences.

The overall compatibility degree $K(LO, P)$ of a learning resource LO to a learner model P is a value in $[0, 1]$ defined in terms of the compatibility between LO and each component of P. Namely, we define:

$$K(LO, P) = \max_{p \in P} K(LO, p)$$

We use the 'max' operator since we express the overall compatibility as disjunction of elementary compatibilities computed between the LO and the single model components. The compatibility degree between a LO and a component p is defined by matching the fuzzy metadata shared by the LO and the component, that is:

$$\begin{aligned} K(LO, p) = AVG\{K(\mu_{LO}, \mu_p)| \\ \exists a \in A \; s.t. \; \langle a, \mu_{LO}\rangle \in LO \wedge \langle a, \mu_p\rangle \in p\} \end{aligned} \tag{5}$$

where AVG is the standard mean, which is used as a particular case of aggregation operator, and $K(\mu_{LO}, \mu_p)$ is the compatibility degree computed between two fuzzy sets. To evaluate the compatibility degree between two fuzzy sets, we adopt the Possibility measure [7], that evaluates the overlapping between fuzzy sets as follows:

$$K(\mu_{LO}, \mu_p) = \max_{x \in Dom(A)} \{\min(\mu_{LO}(x), \mu_p(x))\} \tag{6}$$

The possibility measure evaluates the extent to which there exists at least one common element between two fuzzy sets. This measure is particularly suitable to quantify compatibility between fuzzy metadata, since we assume that two metadata are compatible if they share at least one value of a given attribute.

3 Adapting learner models

In order to derive a model that captures the preferences of a learner during his/her interaction with the e-learning system, we propose a strategy that dynamically creates and updates learner models during time.

For each learner, the proposed adaptation strategy starts with an empty model. During the adaptation process, the model is dynamically updated as the considered learner accesses to the learning resources. The approach used for deriving and updating a learner model resembles to a competitive learning procedure [2], with some variations necessary for dealing with the various components of the learner model.

The adaptation process works as follows. Given a learner, his/her model is initially defined as an empty set, i.e. $P \leftarrow \emptyset$. Next, whenever a learning resource LO_t is accessed by the learner at time t, the model is updated in the following way. For each model component $p \in P$, the compatibility degree between LO_t and p is computed and the model component giving the maximum degree is selected, that is:

$$p^* = \arg\max_{p \in P} K(LO_t, p)$$

If the compatibility degree $K(LO_t, p^*)$ is low (i.e. it is less than a fixed threshold δ_t[1]), it means that there is no compatibility between the learning resource LO_t and the existing model components, hence a new model component is added to P using the same metadata of LO_t, i.e.:

$$P \leftarrow P \cup \{LO_t\}$$

Conversely, if the compatibility degree is high (i.e. $K(LO_t, p^*) \geq \delta_t$), the model component p^* is updated so as to resemble to LO_t. The update concerns all attributes and it is performed according to the following rules. For each $a \in A$, we denote as $\mu^a_{p^*}$ the fuzzy set in metadata $\langle a, \mu^a_{p^*} \rangle \in p^*$ (if the attribute a is not used in the model component, we consider the degenerate fuzzy set, i.e. the fuzzy set such that $\mu^a_{p^*}(x) = 0$ for each $x \in Dom(a)$). Similarly, we define the fuzzy set $\mu^a_{LO_t}$ in metadata $\langle a, \mu^a_{LO_t} \rangle \in LO_t$. The fuzzy set $\mu^a_{p^*}$ is updated as follows:

$$\mu^a_{p^*}(x) \leftarrow (1 - \alpha_t)\mu^a_{p^*}(x) + \alpha_t \mu^a_{LO_t}(x)$$

[1]The value of δ_t serves to establish whether to create a new model component or update an existing one. We observe that for $\delta_t = 0$ no new model component is created, independently on the value of compatibility degrees between LOs and the existing model components. On the other hand, for $\delta_t = 1$, new model components are created for every distinct LO accessed by the learner. In this work, we choose $\delta_t = 0.5$ for $t \leq 0.1N$ and $\delta_t = 0$ for $t > 0.1N$, where N is the estimated number of total accesses a learner makes to the systems. In this way, new model components are generated only in the initial phase whenever incompatible LOs are accessed by the learner.

for each $x \in Dom(a)$. The new fuzzy set μ_{p^*} results a linear combination of its older version and the fuzzy set μ_{LO_t}. We can observe that if $\alpha_t = 0$ no update takes place; on the other hand, if $\alpha_t = 1$ the previous definition of μ_{p^*} is replaced with μ_{LO_t}. The parameter α_t is tuned dynamically so as to favor adaptation during earlier stages of the process. We refer to this early phase as *plastic phase*. As t increases, we make the adaptation less influencing so as to stabilize the model components. We refer to this last phase as *stable phase*. To achieve this behavior, the parameter α_t varies according to the following law:

$$\alpha_t = \exp(-\alpha(t - 1))$$

where the value of α is set empirically so that α_t is greater than 0.5 for $t < 0.1N$, being N the estimated number of total accesses a learner makes to the system. In other words, according to the frequency of learner accesses to the LOs, we estimate that the first 10% of time is used only to create the learner model (LOs are not suggested during this initial stage) while the remaining time is used to update the model as well as to recommend LOs to the learner. This can be achieved by setting:

$$\alpha = \frac{10 \log 2}{N - 10} \approx \frac{7}{N - 10}$$

4 Preliminary simulation results

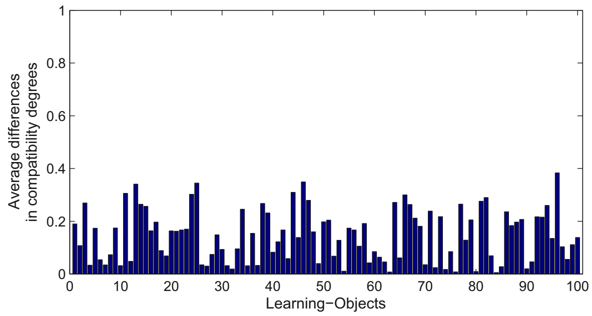
The proposed approach for deriving adaptive learner models was tested in a simulated environment. The simulation was aimed at verifying the ability of our approach in creating several model components that correspond to distinct preferences of a learner.

We randomly generated 100 LOs with uniform distribution. We assumed that each LO was characterized by the presence of five attributes, conventionally named a_1, a_2, a_3, a_4 and a_5. Each attribute had a three-valued domain, i.e. $Dom(a_i) = \{v_1, v_2, v_3\}$. To verify the ability to derive different model components from the same learner, we defined an ideal model made up of three components, as follows:

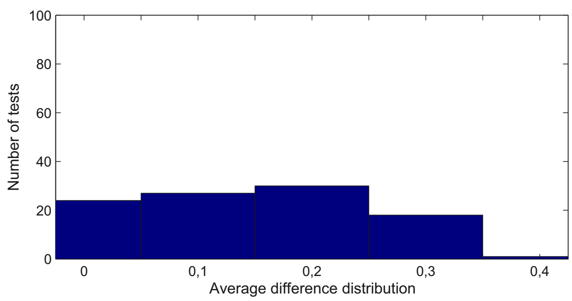
$\langle a_1, \{v_1/1\}\rangle$	$\langle a_2, \{v_2/1\}\rangle$
$\langle a_3, \{v_3/1\}\rangle$	$\langle a_4, \{v_3/1\}\rangle$
$\langle a_2, \{v_2/1\}\rangle$	$\langle a_4, \{v_3/1\}\rangle$

A linguistic interpretation of the model might be the preference for either one of the following types of LOs:

- LO with General Topic "Computer Science" ($\langle a_1, \{v_1/1\}\rangle$) and of "Theoretical" Genre ($\langle a_2, \{v_2/1\}\rangle$);
- LO targeted to "researchers" ($\langle a_3, \{v_3/1\}\rangle$) with Specific Topic on "Programming" ($\langle a_4, \{v_3/1\}\rangle$);



(a)



(b)

Figure 3. Average differences in LO compatibility degrees with ideal and derived model (a) and their distribution (b) over 100 tests.

- LO with Specific Topic on "Programming" ($\langle a_4, \{v_3/1\}\rangle$) of "Theoretical" Genre ($\langle a_2, \{v_2/1\}\rangle$).

In this simulation, we assumed that the learner may make a total number of accesses $N = 50$. To simulate the learner behavior, we generated three probability distributions for the random peek of LOs. The following rule defines the first probability distribution, which is related to the first ideal model component:

$$Prob(LO_j) = \frac{\mu^{a_1}_{LO_j}(v_1) \cdot \mu^{a_2}_{LO_j}(v_2)}{\sum_{h=1}^{100} \left(\mu^{a_1}_{LO_h}(v_1) \cdot \mu^{a_2}_{LO_h}(v_2)\right)} \quad (7)$$

The remaining probability distributions were defined accordingly, using the second and the third component. The simulation proceeded by carrying out the following steps:

1. Select randomly an integer number in $\{1, 2, 3\}$ (uniform distribution) to select a probability distribution for random peek of LOs;

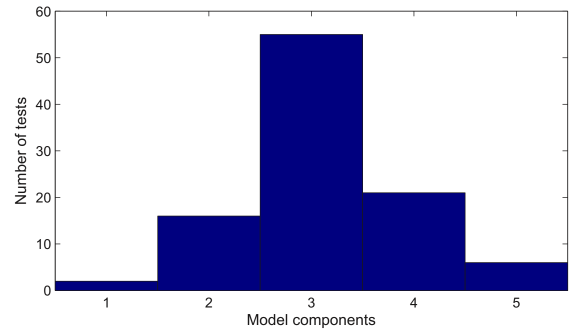


Figure 4. Distribution of the number of model components in 100 tests

2. Select a LO according to the corresponding probability distribution;

3. Update the learner model as described in Section 3. Eventually, a new model component is created if the LO is incompatible with the existing model components.

These three steps were iterated N times. At the end of the adaptation stage, for each learning resource LO_j we compared the compatibility degree of LO_j to the ideal model with the compatibility degree of LO_j to the actual model. We expected that the two compatibility degrees would not differ too much. The entire simulation was run 100 times to gather statistically significant results. Fig. 3(a) shows the average values of the differences between the compatibility degrees of each LO with the ideal and the derived model. In fig. 3(b), the distribution of such values is shown. As it can be observed, about in 50% of trials differences between compatibility degrees were less than 0.15, and this percentage increases to about 75% in correspondence to a difference of 0.2. These results indicate a good performance of the adaptation algorithm, in consideration of the random pick of the learning resources (7) that prevents the derived model to converge exactly to the ideal one.

Also, in fig. 4 we report the distribution of the number of model components generated in 100 tests. It can be seen that in the most frequent case (55 tests) three model components were generated, thus reflecting the structure of the ideal model. In some cases (18 tests) the number of model components was less than required. Obviously, in these cases the matching performance was not fully satisfactory. In the remaining cases, more than three model components were derived. Anyway, this unnecessary redundancy did not hampered the matching performance.

5 Discussion

The last years were characterized by a strong interest in adaptive e-learning systems able to suggest contents to learners by adapting them to their preferences and needs. In such systems, user modeling covers a key role for the definition of models that represent in a significant manner preferences of learners. Another important aspect that has to be considered in the context of user modeling is the ability to dynamically update the derived models so as to adapt them to the constant changes in the interests of the learners when they choose among the learning resources to visit.

In particular, this paper proposed a fuzzy representation of learner models and a strategy for dynamic updating of such models based on procedure that resembles a competitive learning algorithm. The adaptation strategy updates continuously models taking into account the resources that each learner chooses during the interaction with the system. This strategy is essentially characterized by two phases: an initial phase (plastic phase) that is more sensitive to updates and a second phase (stable phase) that is less sensitive to adaptation, thus enabling the suggestion of LOs.

Results obtained by the simulations carried out have shown that the adaptation algorithm converges to significant models including a number of components useful to describe the changing interests of learners.

Future research will investigate on methods for refining the adaptation procedure by taking into account several issues, such as merging similar model components or pruning useless model components.

References

[1] N. Adler and S. Rae. Personalized learning environments: The future of e-learning is learner-centric. *E-learning*, 3:22–24, 2002.

[2] S. C. Ahalt, A. K. Krishnamurthy, P. Chen, and D. E. Melton. Competitive learning algorithms for vector quantization. *Neural Networks*, 3(3):277–290, 1990.

[3] P. Brusilowsky. Adaptive and intelligent technologies for web based education. *Intelligent Systems and Teleteaching*, 4:19–25, 1999.

[4] P. Dolog, N. Henze, W. Nejdl, and M. Sintek. Personalization in distributed elearning environments. In *Proc. of WWW2004 The Thirteen International World Wide Web Conference*, New York, USA, 2004.

[5] V. M. Garca-Barrios. Adaptive e-learning systems: Retrospection, opportunities and challenges. In *Proc. of International Conference on Information Technology Interface (ITI 2006)*, pages 53–58, Cavtat, Croatia, June 12-22 2006.

[6] C. Jing and L. Quan. An adaptive personalized e-learning model. In *Proc. of 2008 IEEE International Symposium on IT in Medicine and Education*, pages 806–810, 2008.

[7] W. Pedrycz and F. Gomide. *An Introduction to Fuzzy Sets. Analysis and Design*. MIT Press, Cambridge (MA), 1998.

[8] L. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

[9] L. Zadeh. Precisiated natural language (PNL). *AI Magazine*, 25(3):74–91, 2004.

[10] L. Zadeh. Is there a need for fuzzy logic? *Information Sciences*, 178:2751–2779, 2008.

Adaptive learning using SCORM compliant resources

Lucia Monacis, Rino Finamore, Maria Sinatra
Department of Psychology
University of Bari
70125 Bari, Italy
lucymonacis@yahoo.it,
finamore.rino@unibari.net

Pierpaolo Di Bitonto, Teresa Roselli, Veronica Rossano
Department of Informatics
University of Bari
70125 Bari, Italy
{dibitonto, roselli, rossano}@di.uniba.it

Abstract — **In recent years great efforts of e-learning research have been focused on customising learning paths according to user preferences. Starting from the consideration that individuals learn best when information is presented in ways that are congruent with their preferred cognitive styles, the authors built an adaptive learning object using the standard SCORM, which dynamically related different learning content to students' cognitive styles. This was performed in order to organize an experimental study aimed at evaluating the effectiveness of an adaptive learning object and the effective congruence of this adaptive learning object with the presentation modes and cognitive styles.**

The sample was made up of 170 students enrolled in two different University courses. The data were gathered by a Cognitive Styles Questionnaire to identify each student cognitive profile, a Computer Attitude Scale to assess the computer-related attitude, and Comprehension Tests. The results indicated that there was a good flexibility of the adaptive learning object, and that analytic and imaginer subjects showed more positive computer attitudes related to a better comprehension of the learning content.

Keywords-component: cognitive style, adaptive learning object, SCORM standard

I. INTRODUCTION

Nowadays we are experiencing a radical change in the didactic and education system which is leading several schools, universities, and companies to adopt state of the art Web based technologies as a new means of managing and sharing knowledge. Such a change is favoured by the numerous advantages guaranteed by Distance Education. One of the most notable and often mentioned benefits is flexibility in time and space: the majority of programs allow students to learn when and where it is more convenient for them, without the grind of the traditional classroom setting. On the other hand, in Distance Education the lack of the teacher's continuous monitoring of the student's activities can cause distraction and frustration.

In the last thirty years, the Adaptive Hypermedia have been the focus of Distance Education research. In [1] Brusilovsky considers the problem of building adaptive hypermedia systems and states that the student's background, experience and preferences should taken into account. As a consequence, in recent years a great number of works have been carried out in the adaptive hypermedia and user modelling research [2, 3, 4, 5].

Moreover, as psychological investigations have revealed that individuals learn best when information is presented in ways that are congruent with their preferred cognitive styles [6], the effort of research in the adaptive learning area has been focused on the use of students' cognitive and learning styles, as reported in [7, 8, 9].

The authors' research work was aimed at defining a technique to design and build adaptive learning paths in e-learning environments using the standard SCORM. In [10] a first technique to adapt the learning content of a SCORM package according to the learner cognitive styles was presented. The Italian Cognitive Styles Questionnaire defined by [11] De Beni, Moè, and Cornoldi was used to define how to tailor the learning content to the students' profiles.

The main issue for defining an effective tailoring technique, is analysis of the relationship between cognitive styles and the way of presenting learning material.

In this context, an experimental study was carried out to assess both the effectiveness of an adaptive learning object which relates different learning content to the students' cognitive styles, and the congruence between the presentation modes and cognitive styles.

II. PSYCHOLOGICAL ISSUES: COGNITIVE STYLES

Since the '80s several studies have shown that the use of Distance Education systems improve the performance of those students who interact with these environments compared to those who interact in a traditional classroom [12, 13, 14]. However, since the '90s many researchers have been consistently asking how the structure and the learning material interact with students' cognitive styles. Previous investigations focused generally on the physical organisation and external appearance of the learning material, i.e. the physical layout, such as the size of the viewing window, the inclusion of headings, etc. [15]. Other studies [16, 17] stated that the manner of presentation as represented by verbal, pictorial or auditory modes affected learning performance according to cognitive style.

As far as the concept of cognitive style is concerned, it should be noted that it refers to the specific way that an individual being codes, organizes, and performs with information, leading to a cognitive management of learning strategies [18]. Consequently, there are several different cognitive styles.

In 1991 Riding [19] suggested that all cognitive styles could be categorised according to two orthogonal dimensions: the wholist-analytic dimension and the verbaliser-imager one.

The former dimension can be considered as the tendency to process information either as an integrated whole or in discrete parts of that whole. Thus, wholists are able to view learning content as complete wholes, but they are unable to separate them into discrete parts; on the contrary, analytics are able to apprehend learning content in parts, but they are unable to integrate such content into complete wholes.

The latter dimension can be considered as the tendency to process information either in words or in images. Verbalisers are better at working with verbal information [20], whereas imagers are better at working with visual and spatial information, i.e., with text-plus picture.

Starting from these introductory statements, our research work aims at defining a technique to build an adaptive learning object standard SCORM which can be tailored on the basis of the learner cognitive styles. The cognitive styles were classified according to the Italian Cognitive Styles Questionnaire that provides the different users' profiles divided into wholists, analytics, verbalisers, and imagers. The Questionnaire details are presented in section V.

III. TECHNOLOGICAL ISSUES: STANDARD SCORM

The standard SCORM (Sharable Content Management Metadata) is one of the most widespread standards used for building LO because it allows the interoperability between the content (LO) and the container (LMS).The standard thus offers the possibility of defining didactic content that can be easily adapted to the users-LMS interaction. In order to understand how user adaptation can be possible, some details on the standard SCORM should be given.

The SCORM consists of: the Content Aggregation Model (CAM), which describes how the SCORM package should be built; the Run Time Environment (RTE), which simulates the LMS behaviour; the Sequencing and Navigation (SN), which describes how each LO component should be aggregated in order to offer different learning paths to the users.

The CAM specification describes the components used in a learning experience, how to package and describe those components and, finally, how to define sequencing rules for the components. Figure 1 depicts the organisation of a learning content in a SCORM package. The learning content is made up of assets, which are the smallest part of an LO, such as a web page, a text or an image. The assets are, then, aggregated in Sharable Content Objects (SCO), which have to be tagged with metadata in order to facilitate their search and reuse. The SCOs are, in fact, the smallest unit that can be lunched and traced by the LMS. The next level is the aggregation, which is not a physical file but just a representation of the organisation of a SCORM package. The aggregation represents the rules of sequencing used to aggregate the different SCOs and/or assets. The SCORM package may, therefore, consist of one or many SCOs and assets.

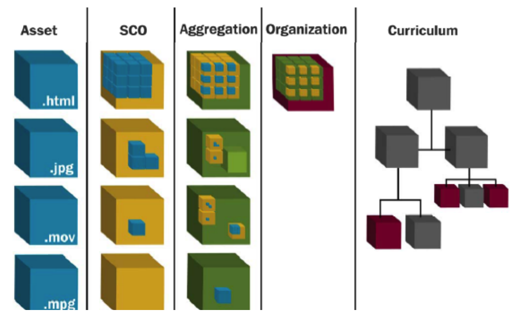


Figure 1. SCORM package organisation [ADL]

At this point the Sequencing and Navigation specifications are used to define the tree structure and sequencing behaviour used to navigate among the different components of the SCORM building different learning paths. Using the SN it is possible, during user interaction, to dynamically choose which SCO has to be launched by the LMS. This allows the LMS to build different customised learning paths in the same SCORM package.

IV. THE DOMAIN CONTEXT

The real context chosen for the experimental study is the course of Psychology of Communication for undergraduate students belonging to different degree courses: Informatics and Digital Communication, and Humanities. The use of the same content in different learning contexts and with different learners (with different backgrounds and different learning approaches) allowed the authors to assess whether learning content customisation, on the basis of the learner preferences, could be successful at any time.

The chosen content dealt with three different topics concerning communication: structure, the various functions and the persuasive models of communication.

Each topic was divided into two didactic units: the first one represented the learning content, described using different presentation modes according to the users' cognitive styles; the second one represented the reinforcement of the same learning content. Moreover, each didactic unit was followed by a multiple-choice test (Comprehensive Test). The overall number of tests was 24. The navigation among the different units will be explained in section VI.

V. THE COGNITIVE STYLES QUESTIONNAIRE

Defining the rules to be implemented in the SN of the SCORM package required the definition of the learner cognitive styles obtained by the submitting the Cognitive

Styles Questionnaire developed in 2003 by De Beni, Moè, and Cornoldi [11].

It consists of two parts and nine items with a 5 points Likert scale for each style. To assess either the wholistic style or the analytical one, students have to observe a figure for thirty seconds and reproduce it. The figure, a sort of Rey's test (1966) revised experimentally by Cornoldi, De Beni, and the MT Group [21], included both a global configuration and some elements regarding a missile, a big pencil, a little flag, single shapes, etc. Nine items were provided to assess the students' preference towards the wholistic style (5 items) or the analytic one (4 items). All items concerned both the analysis of figure (3 items) and various situations (6 items). As regards the students' preference towards verbaliser or imager styles, twelve words and twelve images are proposed. Students have to answer nine items: four items concerned the verbaliser style and five items concerned the imager one. All items referred to the required task consisting of writing the learning material. The questionnaire had to be completed within 25 minutes.

In order to define the user cognitive style a score has to be calculated according to the rules provided in the questionnaire. In both cases the score can vary from 9 to 45. The higher the score obtained, the higher the subject's preference for the wholistic style, in the first case, and for verbaliser style in the second one. Therefore, the questionnaire result for each student is an ordered list of cognitive styles. This ordered list allows the software agent to choose the most appropriate presentation mode for each student. This information is recorded in the learner profile used by the SN rules to select the SCO to be launched.

VI. LEARNING OBJECT ORGANISATION

In order to build an interoperable LO that could be easily integrated into any e-learning environment, the SCORM standard was chosen. In designing a SCORM package, the first issue to consider is the granularity of each individual SCO. Since the first definition of LO was given [22], it is well known that the most difficult problem is the definition of the optimal size of an LO for it to be sharable, reusable and effective. If the LO has a low level of granularity, for example a chemistry course, it would be difficult to reuse without changes in other contexts, such as in an Engineering curriculum. On the other hand if the LO has a high level of granularity, for example an animation of a chemical reaction, it could be reused in many contexts, in different ways and with different learning goals, such as in a lesson aiming at showing the atom metaphor for LO. But, if the LOs are small and have a high level of granularity, it will be impossible for a computer agent to combine them without the intervention of a human instructional designer. This problem, called reusability paradox, has been formalised by Wiley [23]: *if a learning object is useful in a particular context, by definition it cannot be reused in a different context, on the other hand, if a learning object can be reused in many contexts, it is not particularly useful in any*.

In our context, in order to obtain a high level of personalisation of the learning content a high level of granularity was chosen. Therefore, each SCO, the smallest unit traceable by a LMS, contains a didactic unit for a cognitive style (i.e. the Persuasive Models of Communication for imaginer learners) and the Comprehensive Test (CT) useful for verifying the learner information acquisition. The overall number of the SCOs is 24, since we had three topics, for each of them two didactic units represented using four presentation modes.

The SCOs, then, are organised using a tree aggregation form that represents the logical organisation of the learning content given by the domain expert and described in section IV. The Sequencing and Navigation rules are used to explore the tree choosing the right SCO to be launched according to the user's interaction. Figure 2 depicts part of the LO navigation. Each single box represents a SCO, which contains a specific domain concept for a cognitive style (i.e. Pervasive Communication for imaginer learners) and the Comprehensive Test. The arrows show the navigation flow among the SCOs: the ones represented using the straight line indicate that the learner passes the Comprehensive Test, otherwise the dotted line is used. If the learner passes the CT, the new SCO, launched by the LMS, will contain the next learning content using the same cognitive style or presentation mode (i.e. verbaliser in the figure). In the event of a learner responding incorrectly, a reinforcement (using the same cognitive style) will be presented, and, finally, a new CT is presented to the learner. If the learner passes this CT, s/he can go on in the learning path using the same cognitive style. If the learner fails the CT twice, we assume that s/he needs to study the content using a different presentation mode. Thus, the LMS launches the SCO that contains the same learning content depicted using a different cognitive style according to the leaner information profile.

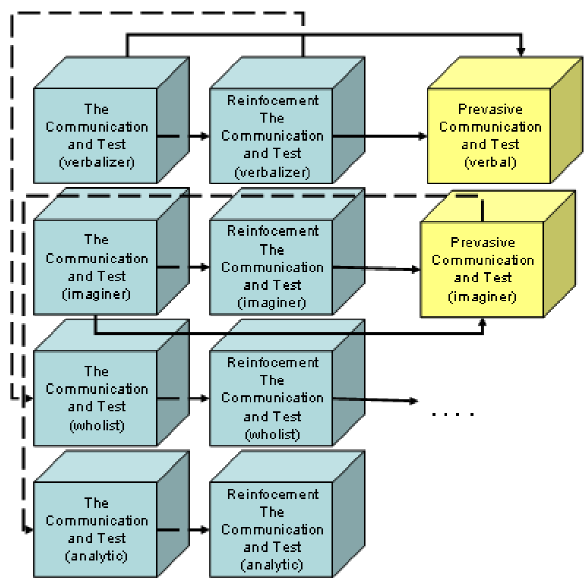


Figure 2. The adaptive learning object organisation

VII. EXPERIMENTAL STUDY

It is important to investigate the relationship between Cognitive Styles, Computer Attitudes, and the manner of presenting learning material in order to assess the effective adaptativity of the Learning Object.

A. Procedure

The general design of our study involved a comparison between students' computer attitudes, their own cognitive style, and the specific learning material. To this purpose, we used assessment tests, preference scales, and the adaptive learning object previously described.

B. Method

1) Participants

A sample of 173 undergraduate students, from both degree courses, was employed for this study. Seven of them were not recorded. The mean age was 20.45 with an SD of 2.03.

2) Instruments

In order to assess subject cognitive styles the questionnaire described in section V was used. The computer attitudes, on the other hand, were assessed using the Computer Attitudes Test (CAS) developed by Al-Khaldi and Al-Jabri [24], reviewed by Shu-Sheng Liaw [25], and translated into Italian, in view of the lack of this kind of scale. Subjects are asked to indicate their perceptions toward computer literacy , liking, usefulness, and intention to use and learn computers. These questionnaires are all seven-point Likert scales (from ''strongly disagree'' to ''strongly agree''), assessing three components of attitudes towards computers, i.e. affective, behavioral and cognitive. The total number of the CAS items is 16, which are divided into 8 items for affective score (1-8 item), 4 items for cognitive score (9,10,15,16) and 4 items for behavioral score (11,12,13,14). The theoretical maximum and minimum possible scores on this scale are 16 and 112 respectively. The attitude towards the computer is assessed according to the following score obtained by students scale:

16-48 low attitude
49-80 average attitude
81-112 high attitude

Moreover, the total score is assessed by the scoring of each subscale, which also determines any predominance amongst them. The "affective" subscale score goes from 1 to 56 and it is divided in the following way:

1-18 low
19-37 average
38-56 high

The "behavioral" and "cognitive" subscale scores go from 1 to 28 and both are divided in the following way:

1-9 low
10-19 average
20-28 high.

C. Results

The data derived from the Cognitive Styles Questionnaire revealed that subjects were labeled as analytics, wholists, verbalisers and imaginers, and in relation to their results obtained from the first Comprehensive Test.

TABLE I.

Cognitive Styles		**Wrong**	**Right**
Wholists	24.70%	56.10%	43.90%
Analytics	26.51%	25.00%	75.00%
Verbalisers	21.69%	61.11%	38.89%
Imaginers	27.11%	28.89%	71.11%
Tot	100.00%	41.57%	58.43%

Data showed that the best scores in the first Comprehensive Test were obtained by analytics and imaginers, 75.00% and 71.11% respectively; on the other hand, the worst ones were verbalisers with 61.11% and wholists with 56.10%. In order to analyze a possible association between 69 wrongers, (amounting to 41.57%) and Computer Attitudes scores, the CAS instrument was used and the data are illustrated in the following table.

TABLE II.

Cognitive Styles	**Computer Attitude Scores- Low**	**Computer Attitude Scores- Average**	**Computer Attitude Scores- high**
Wholists 56,10%	34.78%	43.48%	21.74%
Analytics 25,00%	9.09%	63.64%	27.27%
Verbalisers 61,11%	54.55%	27.27%	18.18%
Imaginers 28,89%	23.08%	23.08%	53.85%

Data indicated a significant effect of Computer Attitude scores on cognitive style. Specifically, verbalisers and wholists obtained a higher score in low attitude, whereas analitycs and imaginers obtained higher score in average-high attitude.

This led to affirm that analytic and imaginer subjects demonstrated more positive attitudes towards the computer than verbaliser and wholistic ones.

2nd STEP

At this point 69 subjects were asked to do the second CT, in order to obtain the academic success. Results are presented in the table below.

TABLE III.

Cognitive Styles		**Wrong**	**Right**
Wholists	33.33%	60.87%	39.13%
Analytics	15.94%	18.18%	81.82%
Verbalisers	31.88%	68.18%	31.82%
Imaginers	18.84%	46.15%	53.85%
Tot	100.00%	53.62%	46.38%

Given the high percentage of mistakes, i.e. 53,62%, this result was analyzed with the Computer Attitude Scale in order to confirm the relationship between cognitive style and Computer Attitude.

TABLE IV.

Cognitive Styles	Computer Attitude Scores- Low	Computer Attitude Scores- Average	Computer Attitude Scores- High
Wholists	35.71%	50.00%	14.29%
Analytics	0.00%	100.00%	0.00%
Verbalisers	66.67%	26.67%	6.67%
Imaginers	16.67%	33.33%	50.00%

Data in table IV show the presence of the significant effect of Computer Attitude and Cognitive Style. Verbaliser and wholist subjects obtained the lowest score in the High Computer Attitude, confirming less positive attitudes towards computer; whereas analytics and imaginers presented the lowest score in the Low Computer Scale.

Moreover, given the high level of wrong responses of wholists and verbalisers to the 2nd Comprehensive Test, the SCORM guided subjects to their second preferred cognitive style and it started up the third step of object Learning Task.

3rd STEP

Data indicated the new percentage of sample distributed according to the switching of the cognitive style.

TABLE V.

Cognitive Styles	Subjects	Wholists	Analytics	Verbalisers	Imaginers
Wholists	37.84%	-	14.29%	28.57%	57.14%
Analytics	5.41%	50.00%	-	50.00%	0.00%
Verbalisers	40.54%	13.33%	53.33%	-	33.33%
Imaginers	16.22%	33.33%	50.00%	16.67%	-

After the 1st Comprehensive test, 78.38% of responses were found to be correct.

TABLE VI.

Cognitive Styles		Wrong	Right
Wholists	37.84%	0.00%	100.00%
Analytics	5.41%	23.08%	76.92%
Verbalisers	40.54%	16.67%	83.33%
Imaginers	16.22%	30.77%	69.23%
Tot		21.62%	78.38%

On the other hand, the presence of some wrong responses, i.e. 21.62%, led to comparison with the CAS.

TABLE VII.

Cognitive Styles	Computer Attitude Scores- Low	Computer Attitude Scores- Average	Computer Attitude Scores- High
Wholists	-	-	-
Analytics	66.67%	33.33%	-
Verbalisers	100.00%	-	-
Imaginers	75.00%	25.00%	-

From the scoring it emerged that nobody obtained any score in the High Computer Attitude and wholists scored nothing in CSA. Furthermore, the verbalisers were found to be the most negative towards computers. In the Average Computer attitude, analytics were more positive than imaginers

VIII. Discussion

In the first CT, subjects with correct responses were 58.43% in relation to their favourite cognitive style. The best results were obtained by subjects with a positive Computer Attitude: more than half of the subjects 31.93% showed a higher attitude towards the computer. Amongst them, analytics were found to be the most positive in object learning, followed by imaginers, wholists and verbalisers.

Moreover, analytics confirmed their positive attitude in behavioral and affective subscales during the two Comprehensive Tests: in the first CT 35% of analytics were behavioral, whereas in the second CT 50% were found to be affective; amongst imaginers 40% of them were “affective” and 60% “behavioral”, respectively in the first and the second CT.

Furthermore, in the switching cognitive style data confirmed that those with correct CT and more positive attitude were analytics with 34.48%, followed by imaginers with 31.03% and wholists and verbalisers, both with 17.24%.

IX. Conclusions and Future Works

The pilot study has demonstrated one of the main advantages of Computer Supported Learning, i.e. the customisation of learning paths according to students’ cognitive style, in order to obtain academic success. From a purely IT point of view, this paper has presented an example of an adaptive LO built using the standard SCORM. The customisation of the learning path in the LO has been first defined by learner favourite cognitive styles resulting from the Cognitive Style Questionnaire developed by De Beni, Moè, and Cornoldi and then analysed from Computer Attitude Scale, in order to explain the main reasons for unsuccessful learning.

Future investigations will involve a sample of students of different educational backgrounds. Moreover, the results obtained will allow to define rules in the expert system presented in [10] to adapt any LO SCORM compliant.

REFERENCES

[1] P. Brusilovsky, “Methods and techniques in adaptive hypermedia,” User Modelling and User-Adapted Interaction, 6(2-3):87 129, 1996.

[2] K. VanLehn, "The behaviour of tutoring systems,” International Journal of Artificial Intelligence in Education, 16 (3), pp. 267-270, 2006.

[3] Michalis Xenos, “Prediction and assessment of student behaviour in open and distance education in computers using Bayesian networks,” Computers & Education, 43, 345–359, 2004.

[4] T. I. Wang, K. H. Tsai, M. C. Lee, and T. K. Chiu, “Personalized Learning Objects Recommendation based on the Semantic-Aware Discovery and the Learner Preference Pattern,” Educational Technology & Society, 10 (3), 84-105, 2007.

[5] C. Bravo, W. R. van Joolingen, and T. de Jong, “Using Co-Lab to build System Dynamics models: Students' actions and on-line tutorial advice,” Computers & Education, In Press, Available online 19 March 2009.

[6] R. E. Riding, M. Grimley, “Cognitive style and learning from multimedia materials in 11-year children,” Br. J. Ed. Tech., vol. 30, January, pp. 43-59, 1999.

[7] A. Calcaterra, A. Antonietti, and J. Underwood, “Cognitive style, hypermedia navigation and learning,” Computers & Education, 44 441–457, 2005.

[8] Huey-Wen Chou, Influences of cognitive style and training method on training effectiveness, Computers & Education, 37, 11–25, 2001.

[9] P. Notargiacomo Mustaro, I. Frango Silveira, “Learning Objects: Adaptive Retrieval through Learning Styles,” Interdisciplinary Journal of Knowledge and Learning Objects, vol. 2, 35-46, 2006.

[10] P. Di Bitonto, T. Roselli, and V. Rossano, “A rules-based system to achieve tailoring of SCORM standard LOs,” International Workshop on Distance Education Technologies (DET 2008), part of the 14th International Conference on Distributed Multimedia Systems, Boston, USA, 4-6 September, 2008.

[11] R. De Beni, A. Moè, and C. Cornoldi, AMOS. Abilità e motivazione allo studio: valutazione e orientamento. Questionario sugli stili cognitivi. Erickson: Trento, 2003.

[12] T. R. H. Cutmore, T. J. Hine, K. J. Maberly, N. M. Langford, and G. Hawgood, “Cognitive and Gender Factors Influencing Navigation in Virtual Environment,” Int. J. Hum. Comp. St., vol. 53, pp. 223-249, 2000.

[13] P.-L. P. Rau, Y.-Y. Choong, and G. Salvendy, “A Cross Cultural Study on Knowledge Representation and Structure in Human Computer Interfaces,” Int. J. Ind. Erg., pp. 117-129, 2004.

[14] M. Workman, “Performance and Perceived Affectiveness in Computer-Based and Computer-Aided Education: Do Cognitive Styles Make A Difference?,” Comp. Hum. Behav., vol. 20, pp. 517-534, 2004.

[15] G. Douglas, and R. J. Riding, “The Effect of Pupil Cognitive Style and Position of Prose Passage Title on Recall,” Ed. Psych., vol. 13, pp. 385-393, 1993.

[16] R. J. Riding, and I. Ashmore, “Verbaliser-imager learning style and children’s recall of information presented in pictorial versus written form,” Ed. Psych., vol. 6, pp. 141-145, 1980;

[17] R. J. Riding, and D. Mathias, “Cognitive Styles and preferred learning mode, reading Webster attainament and cognitive ability in 11 year-old-children,” Ed. Psych., vol. 11, pp. 383-393, 1991.

[18] R. E. Riding, and I. Cheema, “Cognitive Styles: An Overview and Integration,” Ed. Psych., vol. 11, pp. 193-215, 1991.

[19] R. E. Riding, Cognitive style analysis. Birmingham: Learning and Training Technology, 1991.

[20] R. J. Riding, D. Mathias, op. cit. ; R. J. Riding, M. Watts, “The effect of cognitive style on the preferred format of instructional material,” Ed. Psych., vol. 17, pp. 179-183, 1997.

[21] C. Cornoldi, R. De Beni, and Gruppo MT, Imparare a studiare 2. Erickson: Trento, 2001, p. 209.

[22] D. A. Wiley, J. B. South, J. Bassett, L. M. Nelson, L.L. Seawright, T. Peterson, and D.W. Monson, Three common properties of efficient online instructional support systems. The ALN Magazine, 3 (2), http://www.aln.org./alnweb/magazine/Vol3_issue2/wiley.htm, 1999.

[23] D. A. Wiley, “Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy,” in The Instructional use of Learning Objects, D. A. Wiley, joint publication of the Agency for Instructional Technology and the Association for Educational Communications and Technology (Online) http://www.reusability.org/read/.

[24] Shu-Sheng Liaw, “An Internet survey for perceptions of computers and the World Wide Web: relationship, prediction, and difference,” Com. Hum. Beh. 18,17–35, 2002.

[25] M. A. Al-Khaldi, I. M. Al-Jabri, “The relationship of attitudes to computer utilization: new evidence from a developing nation,” Com. Hum. Beh., 14 (1), 23–42, 1998.

Organizing the multimedia content of an m-learning service through Fedora Digital Objects

C. Ardito, R. Lanzilotti
Dipartimento di Informatica, Università di Bari, Italy
{ardito, lanzilotti}@di.uniba.it

Abstract - **In this paper a software infrastructure is presented, developed in the ambit of the CHAT ("Cultural Heritage fruition & e-learning applications of new Advanced (multimodal) Technologies") research project to provide context-sensitive services accessible through thin clients such as cellular phones or PDAs. We propose a client/server architecture. On the server side, software modules implement the algorithms to manage the "dialogue" with the user. The client application displays the multimedia contents sent by the server and captures user interaction and contextual data registered by the mobile device carried by the user and/or received from other sensors in the surrounding environment. The multimedia contents are organized through Digital Objects managed by the Fedora open-source content management software. An additional system component, the Authoring Tool, allows end users lacking computer programming expertise to develop multimedia contents for the different services provided. Exploiting the wide applicability of the infrastructure, we have designed an m-learning service, called Explore!, that supports and enhances middle school students' experience of a visit to an archaeological park.**

Keywords — **m-learning, digital objects, Fedora**

I. Introduction

The CHAT ("Cultural Heritage fruition & e-learning applications of new Advanced (multimodal) Technologies") research project aims at developing a software infrastructure that can provide services accessible through thin clients such as cellular phones or PDAs. Such services must be: a) adaptable to the user's personal preferences, on a choice of interaction modalities; and b) adaptive to the physical-virtual context of the human actor carrying the device. The CHAT project focuses on the development of mobile applications related to two particular domains: e-learning and exploring cultural heritage.

In particular, in this paper we report our experience in designing an m-learning service, called Explore!, that supports and enhances middle school students' experience of a visit to an archaeological park [4]. Explore! adopts a learning technique called excursion-game, whose aim is to help students to acquire historical notions while playing a game on a cell phone [1]. The game is structured as a treasure hunt to be played in groups of 4/5 players, who can freely explore a park and discover its hidden secrets. Each group is given a cell phone and the map. The challenge is to carry out ten missions, displayed on the phone screen one at a time, which require students to walk around and look for the mission targets. If students have difficulties in reaching the mission target place, they can ask the "Oracle" for help. Hints provided by the oracle are formulated to support both gameplay and students' learning of the underlying educational content. Players provide their answer to a mission by typing in the place code on the cell phone. After completing the challenge, the group receives "God's gift": they are given the opportunity to view the 3D reconstruction of the identified places on the phone and visually compare how the places probably once looked with the existing remains.

The multimedia content, e.g. information describing the park, contextual sounds, the game storyboard and some user interface elements of Explore! are stored in Digital Objects (DOs). The DOs are managed by an open-source content management software called Fedora (Flexible Extensible Digital Object Repository Architecture) that provides the tools and interfaces for the creation, "ingest" (a Fedora specific term to indicate the operation of inserting the DO and its own components in the repository), management, and dissemination of the content [5]. In order to define the multimedia content and to ingest each DO that Explore! needs to provide the m-learning service, the Authoring Tool has been developed; through a wizard process, it allows users with no computer programming expertise, e.g. history experts, to easily adapt the service to different archaeological parks.

The paper has the following organization: Section II presents the software infrastructure designed for the

CHAT project, firstly illustrating the server and the client component, then describing Fedora, and finally explaining the Authoring Tool. Section III describes the DOs exploited for the Explore! service. Section IV closes the paper.

II. The Software Infrastructure

The work presented in this paper is part of the CHAT project ("Cultural Heritage fruition & e-learning applications of new Advanced (multimodal) Technologies"), which aims at developing a general-purpose client-server infrastructure for situation-adaptive user assistance. In such architectures, dialogue management is typically based on the integration of independent components that execute specific tasks (sometimes, these components are "out-of-the-box", e.g. components for voice recognition).

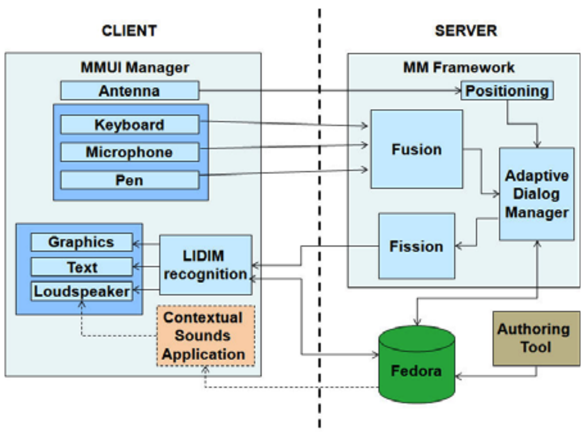


Fig. 1. The CHAT software infrastructure. On the left side the MMUI manager and its own modules running on the clients are depicted. On the right side, the Multimodal Framework represents the server. At the bottom of the figure, the Fedora content management software and the Authoring Tool are shown.

A. *Server components*

The framework proposed in the CHAT project aims at adapting the "dialogue" with the user according to several factors: the service provided, the task currently being executed, the environment in which the user is acting ("context"), the user her/himself and her/his device. These factors are measured and managed by a set of specific software components that together make up the Adaptive Dialogue Manager, as shown in Fig. 1. The *Adaptive Dialogue Manager* is the framework element in charge of a) identifying the most appropriate content to be returned to the client in order to satisfy user's request and b) determining the next system state by updating the models describing the different interaction factors. The Adaptive Dialogue Manager receives its input from a software module, called *Fusion*, which recognizes and combines low-level user input events from different channels (tap or sketch on the screen, voice, gesture, RFID or visual tag scan, etc.) in order to build an overall meaningful input. The output of the Adaptive Dialogue Manager, indicating the most suitable content to be delivered to the user on the basis of the overall interaction state, is refined by a *Fission* module. This retrieves or generates suitable forms of the (possibly multimedia-based) content and takes care of their delivery and synchronization aspects. In particular, it builds an XML-like file, called LIDIM (LInguaggio di Definizione Interfaccia Multimodale, Italian for Language for Defining a Multimodal Interface), which describes the overall interface to be displayed on the client device. The server components, namely Fusion, Positioning, Adaptive Dialogue Manager and Fission, communicate with mobile thin clients through HTTP requests over wireless networks [3].

B. *Client component*

The MultiModal User Inteface (MMUI) Manager running on the client side captures the user's interaction and contextual data registered by the mobile device carried by the user (e.g. GPS position) and/or received from other sensors in the surrounding environment. This information is sent to the server components, i.e. Fusion and Positioning, through HTTP requests. In answer, MMUI receives a LIDIM file that is interpreted by the *LIDIM recognition* module to build the interface and to retrieve the multimedia contents needed, which are managed by the Fedora content management web service.

In the specific case of the Explore! service described in this paper, multiple sounds are emitted during the game execution to ensure a realistic as well as an engaging environment. Sounds are known to fascinate and engage users, particularly younger ones. In virtual reality applications, it has been demonstrated that a sound component added to a virtual scene can help users with navigation, orientation, and the execution of tasks, making the whole experience more enjoyable [1]. Virtual sound sources are placed in various locations in the archaeological park. When the players walk around the park, they will hear the sounds originating from the virtual sources. The perceived sound volume will change according to the distance between the player and each source. The *Contextual Sounds Application* is a software module that defines the user's position and produces a virtual sound environment. It behaves like the JSR234 DistanceAttenuationControl interface available on the

J2ME platform, which controls how the sound originating from a source fades out as the distance from the user grows. The DistanceAttenuationControl is based on the Distance Gain formula; the distance is calculated using the GPS coordinates of the sound sources and the current GPS coordinates of the client (mobile) device.

C. Fedora: multimedia content repository

Fedora (Flexible Extensible Digital Object Repository Architecture) is an open-source content management software that runs as a web service within an Apache Tomcat web server. Fedora provides the tools and interfaces for creation, ingest, management, and dissemination of content stored within a repository, MySQL in the proposed architecture. The use of Fedora has been motivated by the complex nature of the services contents, which can be text, images, audio, videos, etc. Application of different stylesheets to the data and metadata of a Fedora object allows multiple views of the object's content and metadata. Because of this inherent strength and flexibility, it is simple to add new views and data transformations over time as the implementer's and user's requirements change [6]. For example, in Explore! we could choose a smaller level of hints to increase the game difficulty, or we could change the description of the places of interest to satisfy more demanding users. Furthermore, the Fedora architecture is intended to support distributed repositories so it fits our general-purpose client-server infrastructure.

The multimedia contents of the services provided by the CHAT infrastructure are organized in Digital Objects (DOs). A Digital Object in a Fedora repository describes content (data and metadata) and a set of associated behaviours or services that can be applied to that content. Digital Objects comprise the bulk of a repository. A Fedora digital object consists of four parts (Fig. 2):

- *Digital Object Identifier*: a unique, persistent identifier for the digital object.
- *Descriptive Perspective*: the *FOXML metadata* are required by the Fedora repository architecture to facilitate the management of that object. FOXML metadata are distinct from other metadata that are stored in the digital object as content. *Object Properties* describe the object type, its state, the content model to which it subscribes, the dates when the object was created and last modified, and its label. *Relationship Metadata* describe any relationships existing among digital objects in a Fedora repository.
- *Item Perspective*: a *datastream* is the component of a digital object that represents digital content (e.g., digital images, encoded texts, audio recordings). It describes some MIME-typed stream of content: the datastream is a description of this content and a pointer to the content's location. The datastream is not, however, equivalent to a file on a file system because the datastream may encapsulate bytestream content internally in the case of XML content stored with the object, whereas all other content is a reference to content that exists externally to the repository. A datastream can reference any type of content.
- *Service perspective*: a *disseminator* is the component in a digital object that is used to associate behaviours (i.e., services) with the object. All Fedora objects have a default disseminator added at the time of creation so that such services can be immediately retrieved from the repository. The default disseminator allows the repository administrator to get information about the object. Implementers may add any number of customized disseminations to their data objects.

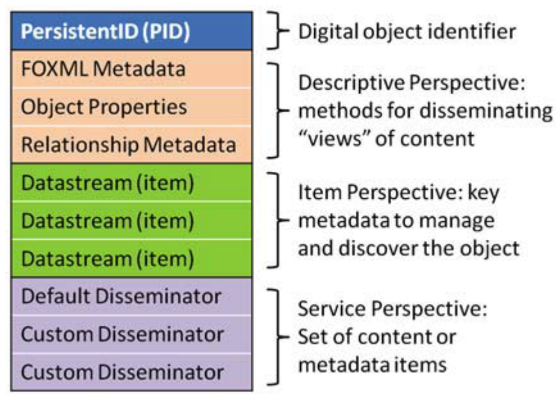


Fig. 2. Fedora Digital Object Architectural View.

In the example in Fig. 3, the image views disseminator allows users to retrieve the content of the object in the views designed by the repository administrators. In this example, a user could retrieve a thumbnail/preview sized image, a pre-defined medium sized image, or a pre-defined high-resolution image. These two last are both generated from a MrSid encoded image, rather than retrieving a static version. The metadata views retrieve the metadata from the object. Users may retrieve the Dublic Core metadata or metadata from an XML type datastream, or both.

D. The Authoring tool

An additional system component, the Authoring Tool, has been developed specifically for the Explore! m-learning service: it allows end users with no computer programming expertise to develop games to be played in different parks. A wizard process guides the user, specifying all the information needed by Explore!: the character to be impersonated (Gaius in the case of Egnathia), the game challenge, the missions, the hints provided by the Oracle. The environment can also be specified: places to be discovered, their photos, 3D reconstructions and GPS coordinates, contextual sounds. Some audiovisual interface elements can be personalized to make them appropriate to a historical period (for example, some icons are suited to a Roman city but not to a Greek city). After the user has terminated the process of specifying the new game, all the information is automatically described by XML files.

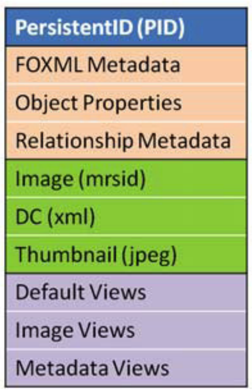


Fig. 3. Fedora Digital Object Image Example.

III. DIGITAL OBJECTS OF THE EXPLORE! SERVICE

The overall information needed to enhance a visit to an archaeological park by exploiting the m-learning service Explore! is stored in Fedora through five digital objects (DOs) described in the following subsections. In particular, the DOs “SITE”, “PLACE” and “CONTEXT” contain the information describing the park, its places of interest and the contextual information (contextual sounds in the specific case presented in this paper). The DOs “SCHEMA” and “INTERFACE” are used to define the game storyboard and some user interface elements. In order to define the multimedia content and to ingest each DO, the CHAT infrastructure supports users through the Authoring Tool that automatically creates an XML file to describe the ingested DO.

A. SITE

The DO “SITE” is identified by the following PID: “site: archaeologicalParkName”. SITE stores the multimedia content about the park: its name, a logo image, and a map. A snippet of the Site.xml file describing the DO ingested for the archaeological park of Egnathia is reported below, as an example:

```
<Site>
  <name>Egnathia </name>
  <logo>
    http://.../fedora/get/site:egnathia/IDLogo
    </logo>
  <map>
    http://.../fedora/get/site:egnathia/IDMap<
    /map>
</Site>
```

The Authoring Tool interface shown in Fig. 4 guides the user in defining the multimedia content about the park and automatically creates the XML file that describes the DO “SITE”.

Fig. 4. Authoring Tool interface for defining and ingesting the DO “SITE”.

B. PLACE

The DO “PLACE” contains the information about places of interest (monuments, roads, market squares, etc) in the park. The multimedia content consists of a text and audio description, photos, GPS position, a map, 3D reconstruction images. Like the other DOs needed for the Explore! service, PLACE is described by an XML file that has the following structure:

```
<place="furnace">
  <description>
    <text>
    http://...:8080/fedora/get/place:furnace/t
    extFurnace</text>
    <audio>
    http://...:8080/fedora/get/place:furnace/A
    udioFurnace</audio>
  </description>
    <photo>
    http://...:8080/fedora/get/place:furnace/P
    hoto_Furnace</photo>
    <map>
    http://...:8080/fedora/get/place:furnace/M
    ap_Furnace</map>
    <gps>
    http://...:8080/fedora/get/place:furnace/g
    ps_Furnace</gps>
    …
```

The user can define the DO "PLACE" through the interface shown in Fig. 5.

Fig. 5. Authoring Tool interface for defining and ingesting the DO "PLACE".

C. CONTEXT

The DO "CONTEXT" contains the information describing the real/virtual environment. In the specific case of the Explore! service it describes each sound source together with the information that the Contextual Sound Application running on the client side needs to produce a virtual sound environment: identification name, GPS position, the maximum distance at which the sound can be perceived and its intensity (in a range between 0 and 100). The sound file to be played is also indicated (see Fig. 6).

Fig. 6. Authoring Tool interface for defining and ingesting the DO "CONTEXT".

D. SCHEMA

While the three previous DOs describe the archaeological park, the DO "SCHEMA" is related to the game to be played. Through the Authoring Tool interface shown in Fig. 7, the name of the game, the instructions (text and audio), and the identification names of the groups that will play are specified.

Fig. 7. Authoring Tool interface for defining and ingesting the DO "SCHEMA".

In order to avoid confusion in the park, each group must carry out the mission in a different sequence. In the example in Fig. 8, the game designer (i.e. the user that defines the game through the Authoring Tool) has set the DO SCHEMA so that the group "Squadra Blu" will face this missions sequence: Bottega Lastricata, Via Traiana, Basilica Civile, Epigrafe, Foro Boario. Thanks to this interface, for each mission the game designer can specify, in text and audio formats, a description of the target, two levels of depth for the hints helping to find it (the "Oracle").

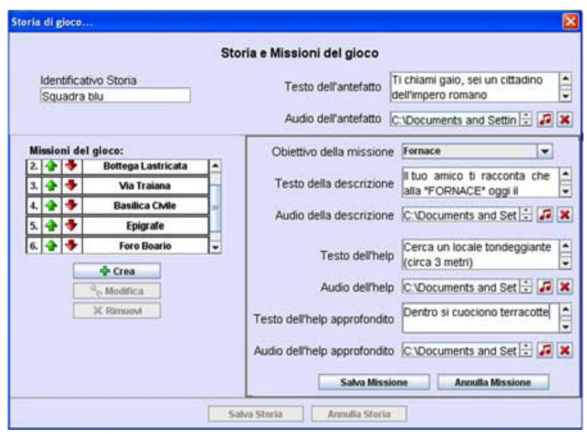


Fig. 8. Authoring Tool interface to personalize the game for each group that will play.

E. INTERFACE

Users can interact with the Explore! service through the interface shown on the client device by the MMUI module. One of the goals of the presented infrastructure and of the use of Fedora DOs is to make the services easily adaptable to new situations, such as another archaeological park, for example. To maintain consistency with the story built around the archaeological park being visited, some interface elements must also be adjustable: for example, the image of a Roman senator announcing a mistake does not fit in with a Greek city.

The DO “INTERFACE” permits the following interface elements to be defined: the images and audio for starting and ending the game, the next mission, and wrong and right answers. The Authoring Tool supports users also in customizing the Explore!’s interface (see Fig. 9).

Fig. 9. Authoring Tool interface for defining and ingesting the DO “INTERFACE”.

IV. Conclusions

In this paper we have presented a software infrastructure that aims at providing general services accessible through thin clients. The point of interest for the DET workshop is that the multimedia content to be provided by the services is organized through Fedora Digital Objects (DOs). Fedora has its own Application Programming Interface (API) to ingest, describe and retrieve the content stored in the DOs. However, preliminary tests we did in the early stage of the CHAT project using this API have demonstrated that the performances decrease with many DOs, which must be all analysed before building an answer to a user’s request. Thus, we have developed an Authoring Tool that has two main advantages: while assists the content designer in defining a DO, it creates an XML file that describes it. In this way, the server finds the information it needs to build the LIDIM file (i.e. the URI to retrieve the content to be displayed on the client interface) in this XML file, thus freeing the Fedora web server from surfing all the DOs. This solution allows us to speed up the overall interaction with the service provided by the infrastructure.

Acknowledgements

The financial support of Italian MIUR through the "CHAT" grant is acknowledged. We thank Professor Maria Francesca Costabile for her valuable guidance in our research.

References

[1] C. Ardito, M.F. Costabile, A. De Angeli, F. Pittarello. Navigation help in 3D Worlds: some empirical evidences on use of sound. *Multimedia Tools and Applications*. vol. 33(2), 2007, pp. 201-216 Springer Netherlands.

[2] C. Ardito, P. Buono, M.F. Costabile, R. Lanzilotti, T. Pederson. Making dead history come alive through mobile game-play. *In Proc. of CHI '07,* ACM Press, 2007, pp. 2249-2254.

[3] C. Ardito, P. Buono, M.F. Costabile, R. Lanzilotti, T. Pederson. Re-experiencing History in Archaeological Parks by Playing a Mobile Augmented Reality Game. *In LNCS 4805*, 2007, pp. 357-366. Berlin, Springer.

[4] C. Ardito, P. Buono, M.F. Costabile, R. Lanzilotti, T. Pederson, A. Piccinno. Experiencing the past through the senses: an m-learning game at archaeological parks. *IEEE MultiMedia*, Oct-Dec 2008. pp. 76-81.

[5] Fedora web-site. http://www.fedora-commons.org/

[6] Fedora tutorial. Tutorial #1: Introduction -- Basic Concepts in Fedora. http://www.fedora-commons.org/resources/

Enhancing online learning through Instructional Design: a model for the development of ID-based authoring tools

Giovanni Adorni[1], Serena Alvino[2], Mauro Coccoli[1]
[1]*Department of Computer, Communication, and Systems Science, University of Genoa*
Viale Causa, 13 - 16145 Genova, Italy
{adorni, mauro.coccoli}@unige.it
[2] *Institute for Educational Technologies – Italian National Research Council*
Via De Marini, 6 - 16149 Genova, Italy
alvino@itd.cnr.it

Abstract

In this paper, a novel point of view in the online learning is provided by the integration of Instructional Design (ID) principles and procedures within the field of Educational Technology. In fact, actual educational technologies and tools do not adequately support teachers when creating, searching for and reusing Learning Objects (LOs); authoring processes are rarely personalized and pedagogical and contextual information is often left aside as well as the implementation of collaborative learning activities. So ID principles and procedures, which normally foster teachers to take the most adequate design choices, can provide a useful support if embedded in the interface of online learning authoring systems and tools. In this respect, Design Models can guide the creation of different types of LOs as well as lesson plans and activities referring to them, through a number of templates and representations. Also, LOs must have a detailed description with pedagogical annotations, in addition to standard metadata, and they should be categorized on the basis of their format so that design of personalized learning paths can be done. According to these premises, this contribution presents a model to develop a new generation of software systems and tools, embedding innovative ID methodologies.

1. Introduction

A number of different definitions and conceptions of "e-learning" can be found in literature. The European Commission defines the e-learning concept as "the use of new multimedia technologies and the Internet to improve the quality of learning by facilitating access to resources and services as well as remote exchanges and collaboration" [1]. On the one side, this definition emphasizes the important role of ICTs and educational technologies, as the way to support the "social dimension" in formal and informal learning processes. Both the increasing success of online learning and CSCL (Computer Supported Collaborative Learning) initiatives [2] and the wide spreading Web 2.0 [3] technologies and social networking tools point out how the "active" and "collaborative" learning is a fundamental paradigm in the current knowledge society. Secondly, the EC definition emphasize the quality of learning, that can be achieved by identifying, sharing and adopting methodologies and best practices for both individual and collaborative learning; to this end, specific Instructional Design (ID) models and strategies have been developed in the last years to support the design of effective e-learning initiatives. The ID is "a construct referring to the principles and procedures by which instructional materials, lessons and whole systems can be developed in a consistent and reliable fashion" [4]). ID principles and procedures are normally rendered explicit through design models (DMs) which are a kind of abstract design rules for a given educational theory or didactic strategy that tells how to organize appropriate materials, lessons or learning scenarios to achieve specific learning objectives. Recent approaches to ID point out that the design process, as it is really put into practice by expert designers, is not a procedure but a problem solving process, guided by heuristics and best practices held as effective for a specific problematic situation [5]. According to this perspective, as demonstrated by a number of studies [6, 7], the alternative of rendering explicit and formalizing heuristics and best practices through DMs for the design and the management of learning resources and activities become more and more relevant in the educational research field. This prospect has become especially significant for the field of CSCL where best practices on how to structure effective individual or collaborative learning process are till now hardly shared by experts [2]. In addition, current trends in the e-learning field [8] are also showing the benefits coming from the investment in the creation, sharing and reuse of Learning Objects (LOs), defined by Wiley [9] as "any digital resource that can be reused to support learning"; this wide accepted definition refers to both standard-based LOs (e.g. SCORM [10]) and LOs supporting collaborative learning [8]. So, at present, the "community" is playing more and more a key-role in the e-learning field both when involved in formal or informal collaborative learning and when involved in the sharing of best practices, through the formalization and the reuse of LOs [11] and DMs [12, 13]. In fact, designers and teachers can create resources and share them within a professional community; these resources can support both individual and collaborative learning and can also be compliant with some international standard so to be interoperable and automatically interpretable by LMS. But how many teachers are able to design an effective LO and describe it adequately so to foster an easy retrieval within a repository? How many of them are able to integrate these resources in active and/or collaborative learning activities? Recent studies [14, 5] have pointed out that e-learning practitioners, especially when unskilled, need to be supported both when creating and describing a LO and when designing a collaborative learning activity.

In this perspective, in line with recent research studies about ID and Leaning Design [12, 13], this contribution aims to present a model to develop a new generation of software systems and tools, which embed innovative ID methodologies; these tools would be able to support unskilled teachers and designers when creating learning materials or designing activities, lessons and courses, so to effectively structure the contents and the activities according to specific heuristics and good practices.

2. ID models and pedagogical metadata: new challenges for online learning authoring tools

Practitioners and researchers, having different educational and technological backgrounds, hardly share a common view on how to support an effective learning process by means of technologies and distance education good practices. The actual added value in the design of applications supporting online learning is the integration of different points of view, raised up by both designers and end-users, and including both educational and technological perspectives. Given the actual situation and the existing platforms and tools, one can observe that many actors are involved in the teaching and learning process and that different objectives are to be considered, depending on specific points of view. Sometimes, some of the objectives are opposite to other ones. International initiatives such as IMS (the Global Learning Consortium), OKI (the Open Knowledge Initiative), and ELF (the E-Learning Framework) [15], put in evidence that different online learning applications may need different characteristics from both a technical and functional point of view: school, universities, industry, corporate, life-long learning have very different requirements. As a matter of fact, there is the convergence of a number of different users' needs in just one system, performing multiple functions and managing different users' roles. Also there is the convergence of many theoretical models and many possible technical solutions. Yet the technological support is not so flexible. The Learning Management System (LMS) has been the main actor of Internet-based education for the past two decades and the main delivery systems for standard compliant LOs. However, the traditional conception of LMS is falling to keep pace with recent advances in education, information and communication technology, and the semantic web [16]. There is much more; thus, a modular architecture is needed for LMSs which can interact with a wide variety of services and tools that may be needed, and even may be different from case to case, for achieving the best results in learning and teaching [17]. In this respect, many researchers have already investigated on how to bridge ID and learning content [18]. In such a context, it is clear that a new generation of software tools designed to simplify the work of the users within their design, teaching and learning activity is needed [19] so that online learning can be significantly improved.

2.1. Embedding design models into online learning authoring tools

Teachers are often unskilled in creating or retrieving educational resources which fit in with the needs of their educational context and often lack competencies on how to share them to foster reusability within a community [14, 20]. In addition sharing educational resources is not a straightforward task for teachers, but requires them a good amount of work both to integrate in their own lessons other people’s productions and to prepare new contributions in easily re-usable and adaptable form; as a consequence, LOs technology struggles to gain momentum and acceptance in the communities of teachers and instructional designers [6]. ID heuristics and practices can provide a fundamental support to teachers for: a) identifying the main constraints characterizing the specific educational context; b) designing effective LOs and learning activities taking into account those constraints; c) searching for reusable resources which can be effectively integrated in a specific learning path. These heuristics are especially important in the context of

CSCL where good practices about how to structure computer-mediated interactions are till now hardly shared by experts [2].

Traditional ID methods and online learning ones can be mould and structured into design models (DMs), i.e. schemata, scripts, meta-models, embedding a specific pedagogical approach, that support teachers in developing educational proposals; these resources can be reused in different educational contexts [6]. In particular, the bridging between collaborative learning and traditional ID methods [12] by means of CSCL scripts, and specially *macro-scripts*, has recently raised a lot of attention. CSCL macro-scripts are models that formalize and represent a sequence of activities aimed at fostering a meaningful learning process in a group [ibid.]. They can be reused and instantiated (adapted, contextualized) in different educational contexts, being formalized at different levels of abstraction; the more abstract level is independent from the content and, generally, it represents the solution to a recurrent educational problem (e.g. Pedagogical Design Patterns); other macro-scripts represent a particular instantiation of the general educational problem, suggesting contents, roles, tools, services, etc., needed to support the activity (e.g. lesson plans and IMS-LD Units of Learning [13]). CSCL macro-scripts can be formalized through different templates and representations, shared and reused, exactly as designers and teachers usually do with LOs.

An innovative approach for the development of a new generation of authoring tools fostering the design of online learning is to integrate DMs in the system interface, in order to support unskilled teachers in the design phase of LOs, activities and modules. Currently, the new research lines focused on the formalization of macro-scripts are systematically translated into practice only by initiatives which implement Learning Design-based [13] authoring tools and platforms (such as *RELOAD* - http://www.reload.ac.uk, *ReCourse* - http://www.tencompetence.org/ldauthor, *LAMS* - http://www.lamsinternational.com or *COLLAGE* - http://gsic.tel.uva.es/collage). Unfortunately, Learning Design theories [21], which propose to represent the learning process by means of formal languages (EML - Educational Modeling Languages), have shown their limits; in fact, although different research lines are currently engaged in identifying methodologies and tools for bringing the Learning Design closer to designers' and teachers' daily practice, technical specifications such as IMS-LD [13] are not so wide-spread in the e-learning field, yet. This is due, from the one side, to their complexity and, from the other side, to the limits embedded in their semantic, which doesn't allow the direct representation of groups and their structuring in collaborative activities [22]. On the contrary, other initiatives demonstrated that macro-scripts, when embedded in the interface of design tools (see e.g. *COLLAGE* [ibid.]), could provide an effective support in the design process [6]. Finally, current trends are pointing out the effective role of diagram-based graphical representation of ID best practices and macro-scripts when embedded in the interface of learning design authoring tools [7]. In addition, some advances have been done in the perspective of modeling the design process of LOs. In the last years some research initiatives tried to define taxonomies of LOs, according to their main technical characteristics and to their semantic dimension [9, 11]. So, different approaches to the design process of LOs have been proposed in literature, some of them integrated in specific authoring tools (e.g. *RELOAD*). But, from an educational point of view, like any other instructional technology, LOs must embed specific ID strategies [9]. So, new approaches overcome the limitations introduced by the main technical specifications [10]: some of these initiatives [23], are now trying to classify LOs according their educational features, such as the embedded didactic strategy. In this perspective it could be feasible to model the structure of different LOs typologies according to their pedagogical approach and to model DMs' structure and flow through text and diagrams. Guidelines for supporting the creation of LOs and for instantiating DMs in a specific context can also be defined. All the ID models, best practices and heuristics involved in the design of this new generation of authoring tools for online learning should be framed in an Instructional Design Reference Model: this model, which constitutes one of the peculiarities of this innovative approach, will guide the design, the development and the integration of the software application by defining a methodological and pedagogical framework for: a) the definition of the main design steps; b) the modeling of the main LO typologies and of a set of reference didactic strategies and DMs; c) the definition of guidelines for LO creation and DMs instantiation.

2.2. Fostering effective reuse through the pedagogical annotation of LOs and design models

Another crucial issue for teachers and designers who want to search for and share educational resources is the identification of proper metadata models or specifications allowing for an effective description and an easy retrieval. Such descriptors should enable users in seeking resources not only on the basis of technical and bibliographic information, but also on the basis of their contextual and educational features. As a matter of fact, the description of the educational needs that inspired the design of a LO, of the underlying assumptions on learning and of the epistemological and pedagogical approaches to the content significantly supports the retrieval of potentially re-usable products and fosters the reflection on their adaptability to the specific context [6, 24]. Such pedagogical metadata sets, together with a user-friendly interface for LOs annotation and retrieval, could support users' motivation to invest their time and efforts in the design, implementation and diffusion of reusable LOs.

A number of metadata specifications has been proposed by various international initiatives (such as LOM [11], EdNA - http://ww.edna.edu.au, TLF http://www.thelearningfederation.edu.au, GEM http://thegetaway.org,). But, the expressive power of these metadata sets is often unsatisfactory with respect to the underlying educational paradigm. In addition, as we pointed out before, teachers, in their practice, usually take advantage not only of learning material directed to students, but also of DMs that represent suggestions, work plans, best practices, etc., developed by their peers. So teachers can be enhanced by repositories that support the description and the retrieval also of this kind of resources. Some international proposals have been presented to improve this situation, such as the POEM model (Pedagogy Oriented Educational Metadata model) [6]; by means of pedagogical vocabularies, validated by different typologies of end-users, this innovative LOM [11] application profile helps designers and teachers to efficiently search for both LOs and DMs.

The challenge of new authoring-tools supporting the creation of both LOs and DMs, such as macro-scripts, is to provide also a support for their description. Contextual and pedagogical annotations can be added to standard metadata to foster an effective sharing of LOs and DMs in suited repositories.

3. The intersection between ID and educational technologies: a multifaceted model for the design of online learning authoring systems

According to the scenario depicted above, new paradigms and methodologies should be considered in the phase of design and development of integrated platforms, systems, and tools in the direction of achieving flexible and open infrastructures, supporting the ID and the development of both multimedia standardized LOs and DMs. Novel functionalities that should be made available in platform and tools for online learning are described in this paragraph, with the aim of demonstrating the potentialities included in considering the intersection between ID and the educational technology. The next generation of online learning services and systems will have to embed ID strategies, methodologies and best practices and will have to be provided with advanced capabilities for information retrieval, indexing and metadata. In particular, such a novel flexible and open infrastructure will have to be able to support online learning designers and teachers in the design of highly customizable courses, materials and activities, based on specific didactic strategies or design models.

A multi-faceted model is proposed (Figure 1) for the definition of the resources to be designed, where structure, LOs, and DMs are kept separate and are to be managed in appropriate environments. Based on the semantic web stack model and on the rdf [25] language as well as on micro-formats [26], also the definition of a conceptual data structure relevant to assets (and their categorization or description, tags, folksonomies, or metadata) should be done.

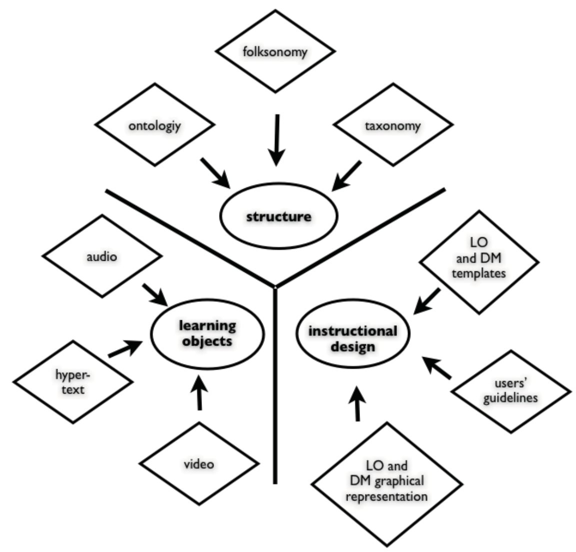


Figure 1: A multifaceted model for the design of online learning authoring systems

According to such a multi-faceted model, the design process of an online learning authoring system should constantly refer to an Instructional Design Reference Model, providing users with a unique pedagogical framework for all the functions supported by the system. It should define and relate each other, well-known ID methodologies and best practices for individual and collaborative online learning, which could drive both the LOs structuring and the DMs instantiation process. To this end, the ID Reference Model should also define templates for the textual and graphical representation of didactic strategies and DMs, which should be made available in libraries, within the system interface. In addition, it also should provide guidelines for supporting the creation of a number of LOs typologies and the instantiation of abstract DMs (e.g. pedagogical patterns) into more contextualized ones (e.g. lesson plans).

The Instructional Design Reference Model should also define the main design steps that will be supported by the system. The ID process can be carried out freely, interacting with external tools and web services, according to one's preference and outside of strict schemata. Design steps can be performed independently or following a suggested design flow.

Users can exploit existing concept maps, i.e. a logical and well-formed diagram clearly expressing the relationships between concepts the LO deals with, or create new ones, The use of such a description is a good method for knowledge representation and allows the link of this encoded knowledge with the relevant resources. Concepts represented in the concept map can be linearized so to easily define the sequencing inside the LO. Due to its intrinsically hierarchical nature, it can be simply turned into a data structure (such as a XML tree) following the above priority rules, thus facilitating the software elaboration and integration while increasing human readability, and exploiting hypertext links. Different types of LOs can be created according to the best suited didactic strategy (changing from case to case). In this perspective, the user would be able to choose among specific templates embedding different strategies and then fill them in with text or multimedia assets. A proper connection to retrieval services and to existing repositories would allow finding assets, LOs and DMs according to the metadata and the tags available. The contextual and pedagogical annotation of LOs and DMs would provide users with a clear pedagogical picture of these resources.

The system could also support designers and teachers to structure DMs, such as lesson plans, selecting a more abstract model (e.g. macro-scripts such as pedagogical patterns) and instantiating it. On the basis of pre-defined templates and diagram-based representations embedding ID best practices, the user would be able to design the structure and the flow of individual and collaborative online activities, lessons or courses; he/she could refer to one or more LOs (linked from the Internet or from a repository) and identify LMS’s features that could support effectively the activity; when designing a collaborative activity, the user could also specify some useful communication tools, roles involved in the learning community and some characteristics of groups engaged in the activity [6]. With the same logical structure used for LOs, contextualized DMs, such as lesson plans, could frame different LOs within a specific activity or course, keeping different data at different levels of representation. Once created LOs or DMs, the user would also need to add metadata, specifying both standard information and contextual or educational features characterizing the resource. System semantic functionalities could support both user input and automated data extraction, thus fostering the sharing and the reuse even when users are unskilled in metadata connotation. Semantic searches, manual resource tagging and metadata extraction should be performed according to a reference application profile which includes specific sets of contextual and pedagogical descriptors. Finally, users can also search for different types of LOs according to the same reference application profile and to custom tags, useful for the folksonomies management in social networking applications and environments [27].

4. Conclusions and future perspectives

The aim of this paper is to prospect an innovative approach for enhancing online learning capabilities, by merging the results of different research fields. On the one hand, this can be obtained by modeling templates and diagrams representing ID strategies, methodologies and best practices so to support their integration in innovative LOs and DMs authoring tools. In such a way unskilled teachers and designers can be fostered in the instructional design of LOs, activities and modules for individual and collaborative learning. On the other hand, users can also be supported in other important ID steps such as the description and the effective retrieval of reusable resources. Specific application profiles should be implemented fitting with designers' and teachers' expressivity needs when interacting with educational software tools and when describing and searching for LOs and DMs.

An ID Reference Model should guide the design and the development of the resulting new generation of authoring systems, which should effectively support unskilled practitioners to: a) organize concepts into LOs according to specific didactic strategies; b) search for assets, LOs and macro-scripts according to different characteristics enhanced by text analysis; c) create DMs, such as lesson plans, framing one or more LOs and embedding specific didactic strategies; d) describe and share their own LOs or DMs in a repository, supported by user-friendly and semi-automatic features, allowing for semantic-based metadata extraction and keeping compliance to the main international standards [10, 11], if required.

From the point of view of knowledge representation, the need is for the definition of description languages and data interchange formats for the management of LOs and DMs, based on existing and emerging international standards. Conceptual representation, topic maps, and ontologies have to be considered among the many possible means for structuring knowledge and different models written with xml, xtm (xml for topic maps), and rdf will have to be made interoperable while seeking for a common equivalent model and efficient transformation rules.

Future work will envisage the design and the development of a first prototype of such authoring tool. Use cases and users scenarios will guide the design phase and a flexible fast-prototyping strategy for software development will be adopted; a number of validation and testing activities will inform different recursive design steps, in a problem-based approach to the design process [5].

References

[1] European Commission, "The e-learning Action Plan: Designing tomorrow's education", available at http://eur-lex.europa.eu/en/com/cnc/2001/com2001_0172en01.pdf, 2001.

[2] T.D. Koschmann, "CSCL: Theory and Practice of an Emerging Paradigm", Lawrence Erlbaum Associates, New Jersey, 1996.

[3] T. O'Reilly, "What Is Web 2.0" available at http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html, 2005.

[4] C.M. Reigeluth, M. Molenda, and L/M. Nelson, "Instructional Design", in L. Nadel (Ed.), Encyclopedia of Cognitive Science, vol. 2, 2003, pp. 574-578.

[5] K.H. Silber, "A Principle-based Model of Instructional Design: A New Way of Thinking About and Teaching ID", Educational Technology, vol.47 n.5, 2007, pp. 5-19.

[6] S. Alvino, P. Forcheri, M.G. Ierardi, and L. Sarti, "A general and flexible model for the pedagogical description of learning objects", in Proceedings of WCC2008, Milan 7-10 September 2008.

[7] S. Alvino, J.I. Asensio-Perez, Y. Dimitriadis, D. Hernandez-Leo. "Supporting the reuse of effective CSCL learning designs through social structures representations", Distance education – Special Issue on Learning Design, 30 (2), In press.

[8] A. Littlejohn, *Reusing online resources*, Kogan Page, London, 2004.

[9] D. Wiley, "The Instructional Use of Learning Objects", available at http://reusability.org/read/, 2000.

[10] IEEE - Learning Technology Standards Observatory, "Scorm RunTime Summary - Introduction", available at http://www.cenltso.org/Users/main.aspx?put=497, 2006.

[11] IEEE - Learning Technology Standards Observatory, "Learning Object Metadata", available at http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf, 2002.

[12] P. Dillenbourg, "Over-scripting CSCL: The risks of blending collaborative learning with instructional design" in *Three worlds of CSCL. Can we support CSCL,* P.A. Kirschner (Ed), Nederland, 2002, pp. 61-91.

[13] IMS, "Learning Design v 1.0 Final Specification", available at; http://www.imsglobal.org/learningdesign/index.cfm, 2003.

[14] S. Alvino, J.I. Asensio-Perez, Y. Dimitriadis, and D. Hernandez-Leo, "Social Structures Representations as aid for effective creation and reuse of CSCL scripts according to a problem-solving approach to ID", in Proceedings of ICALT2008, Santander 1-4 July 2008.

[15] S. Wilson, K. Blinco, D. Rehak, "ELF - An e-Learning Framework - A Summary", DEST-JISC-CETIS, 2004.

[16] D. Dagger, A. O'Connor, S. Lawless S, E. Walsh, and V. Wade, "Service-Oriented E-Learning Platforms: From Monolithic Systems to Flexible Services", IEEE Internet Computing, Vol. 1-3, 2007, pp. 28-37.

[17] C. Chih-Ping, C. Ching-Pao, Y.Chung-Wei, Y. Yu-Fang, "A Web-Service Oriented Framework for building SCORM Compatible Learning Management Systems" in Proc. of ITCC 2004 International Conference on Information Technology: Coding and Computing, Vol.1, 5-7, 2004, pp. 156-161.

[18] C. Knight, D. Gaševic, G. Richards, "An Ontology-Based Framework for Bridging Learning Design and Learning Content", Educational Technology & Society, 9 (1), 2006, pp. 23-37.

[19] M.J. Rosemberg, *Beyond E-Learning*, Pfeiffer, San Francisco, CA, 2006.

[20] T. Hand, "Learning Objects: User Perspectives on the Conditions Surrounding Their Use" in Proc. of World Conference on Educational Multimedia, Hypermedia and Telecommunications, L. Cantoni & C. McLoughlin (Eds.), Chesapeake, 2004.

[21] E.J.R. Koper, "Modelling Units of Study from a Pedagogical Perspective: the pedagogical metamodel behind EML", Heerlen: Open Universiteit Nederland, available at http://eml.ou.nl/introduction/docs/ped-metamodel.pdf, 2001.

[22] D. Hernández-Leo, E.D. Villasclaras-Fernández, J.I. Asensio-Pérez, Y. Dimitriadis, I.M. Jorrín-Abellán, I. Ruiz-Requies, and B. Rubia-Avi, "COLLAGE: A collaborative Learning Design editor based on patterns" in Educational Technology & Society, 9 (1), 2006, pp. 58-7.

[23] A. Ip and I. Morrison, "Learning objects in different pedagogical paradigms", in Meeting at the Cross-roads. Proc. of the 18th Annual Conference of the Australian Society for Computers in Learning in Tertiary Education, G. Kennedy, M. Keppell, C. McNaught, and T. Petrovic (Eds.), Melbourne, 2002, pp. 289-298.

[24] E. Busetti, G. Dettori, P. Forcheri, and M.G. Ierardi, "A pedagogical approach to the design of learning objects for complex domains", Int. Journal of Distant Education Technologies, vol. 5, no. 2, 2007, pp.1-17.

[25] Resource Description Framework - RDF, available at http://www.w3.org/RDF/

[26] R. Khare, "Microformats: the next (small) thing on the semantic Web?", in IEEE Internet Computing, vol. 10 (1), 2006, pp. 68-75.

[27] H.S. Al-Khalifa and H.C. Davis, "Replacing the Monolithic LOM: A Folksonomic Approach", in Proc. of The 7th IEEE International Conference on Advanced Learning Technologies, Niigata, Japan, July 18-20, 2007.

Learning Objects Design for a Databases Course

Carlo DELL'AQUILA, Francesco DI TRIA, Ezio LEFONS, and Filippo TANGORRA
Dipartimento di Informatica
Università di Bari
via Orabona 4, I-70125, Bari, ITALY
{dellaquila, lefons, tangorra, francescoditria}@di.uniba.it

Abstract — Face-to-face teaching involves both theoretical and laboratory lessons. Here, we present an on-line course, developed by an open source software, in order to support traditional face-to-face lessons in the Databases field. The courseware is based on learning objects designed to support different teaching needs derived not only from the Computer Science degree, but also from other university degrees that require database skills. In detail, the courseware is composed of both traditional theoretical lessons, about the Databases theory, and laboratory lessons, about the three most popular DBMSs for relational databases. There are also illustrated some preliminary statistical analyses that we have carried out on the fruition of our databases course from the students of the University of Bari.

Keywords-*Learning Management System, Learning Object, Courseware, Databases.*

I. Introduction

Databases course is a fundamental teaching included in every Computer Science degree course since ever. Nowadays, the exploitation of databases has enormously increased as a result of the spread of modern office automation tools. This trend has led to the introduction of the databases teaching not only in many degree courses of several university Faculties but also in the training of the employees of business companies [1-3]. Of course, each of these courses focuses on different aspects and provides specific skills.

For example, in Computer Science curricula, the traditional database contents have been extended with advanced concepts, in order to fit the course to the current standards and technologies. In fact, with the advent of the World Wide Web and the distributed computing, there is a strong incentive towards the integration of traditional relational databases and XML files [4]. Moreover, new data structures and data definition/manipulation languages have been introduced in order to allow developers to model complex domains and to perform advanced data analyses. These kinds of issues are the topics of *deductive databases*, based on the first-order logic [5], and *object-oriented databases*, based on the object-oriented paradigm [6]. On the other hand, many courses provide only the basic concepts on databases field and adopt simple tools as training environment. For example, it is usually so in the courses for the European Computer Driving Licence (ECDL) [7].

The traditional face-to-face courses must also be supported by didactic material that permit the students to find the main concepts explained by teachers and to affront individually the Databases exam preparation in effective way. Since this teaching field is very wide, it is a hard work to recommend a textbook covering in exhaustive way all the fundamental topics and that meets the needs of the teacher. Moreover, modern teaching strategies drive teachers to provide didactic contents tailored for a specific learning context and encourage students to undertake formative paths according to their own predispositions.

In order to improve the quality of the teaching and the learning processes, e-learning environments represent a valid solution, because they provide high flexibility and reusability of didactic contents, against an accurate phase of design and implementation of the basic learning objects [8].

In this paper, we present our experience about the design of a set of learning objects for teaching a Databases course for the Computer Science students and the related benefits obtained by furnishing didactic contents via an e-learning platform. This courseware is composed of both theoretical and laboratory lessons, and its own learning objects can be easily re-arranged in order to satisfy other teaching needs.

This paper is organized as follows. Section II introduces the e-learning platform used to distribute the didactic contents. Section III shows how the learning objects compose the macro-components of the courseware and explains the design of the learning objects. Section IV reports our personal experience deriving from the adoption of an e-learning platform. Section V introduces the criteria adopted to evaluate the usability of the courseware. Finally, Section VI reports our conclusions.

II. Learning Content Management System

The courseware is hosted on ATutor, that is an open-source e-learning platform, developed at the University of Toronto [9]. The main characteristics of ATutor are the *adaptability* and the *accessibility*.

The term *adaptive* stands for the capability of fitting to a wide range of possible scenarios of teaching and learning. In fact, the e-learning platform defines three roles for registered users: (a) administrator, (b) teacher, and (c) student. Each of these kinds of user can modify the system preferences, according to different levels of privileges. For example, the administrator can add further language packages to the system, while the teacher can define the basic settings for its own courses and also decide the access policies. In fact, some courses need teacher approval for the access to didactic contents. Finally, the student can only customize the user interface and add/remove the visibility of features.

The term *accessible* means that ATutor has been designed in order to allow all kinds of users to perform an easy access to didactic contents, according to the technology adopted for the teaching or the learning process. For example, the e-learning platform is accessible to both users with disabilities, that are assisted by means of particular technologies, and users that do not benefit of modern and advanced technologies, as recent browsers and fast internet.

Further features comprise:

- a web-based interaction, allowing access according to a temporal and spatial independence;
- good levels of security and password management;
- asynchronous and synchronous communication tools, as chat, forum, blog, and embedded email system;
- functionalities for collaborative works;
- traceability of the students activity.

Finally, ATutor is compliant with all the e-learning standards, as IMS [10] and SCORM [11], allowing interoperability among platforms by the import/export of content packages. Moreover, the compliance with the specifications of the W3C XHTML [12] ensures the correctness of the layout according to a device-independence.

III. Structure of the Courseware

The courseware of Databases is composed of two main blocks: (a) Theoretical lessons, and (b) Laboratory lessons (see, Figure 1).

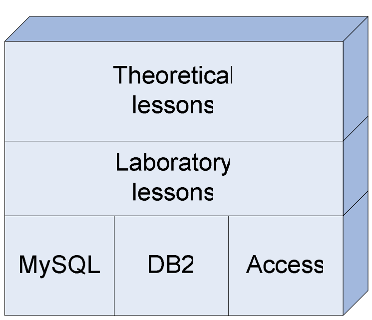


Figure 1. Courseware macro-components.

On turn, the section containing the laboratory lessons is divided into three sub-blocks, each of them describes and explains a specific software tool, the so-called Database Management System (DBMS). A DBMS is used as a reference environment for the lessons held in didactic laboratories. The chosen DBMSs are: MySQL [13], DB2 [14], and Access [15], because they are the most suitable systems to use for training purpose at different levels of difficulty.

The granularity chosen is quite high. In fact, each block of Figure 1 is composed of a set of Learning Objects (LO) that can be easily re-assembled in order to create a new courseware, according to different needs and training targets. The conceptual design of the learning objects is illustrated in the next Sub-sections.

A. Theoretical Lessons

The Databases courseware presents digital contents, in Italian, corresponding to didactic units of the course held in classroom. In addition, the basic concepts of theoretical lectures have been reinforced with exercises that allow the student to improve its preparation for the final examination [16].

The contents of the lectures cover topics according to the suggestions of ACM/CS guidelines [17] for Database courses and following the popular textbook by R. Elmasri and S. B. Navathe [18].

The learning objects that compose the courseware have been organized according to the conceptual map shown in Figure 2 and the teaching approach consists of following the methodological choice to apply the same learning strategy to the lectures in classroom and the course. So, the student can retrieve in the course the same guidelines of the presentations used by the teacher in the lessons.

The course is accessible by using a general index, *i.e.* the navigation tree, that represents the learning objects as a hierarchical tree, such that to avoid students to get lost in hyperspace [19]. In addition, the navigation among the learning objects is organized by means of opportune links that drive the students through a logical path, composed also of exercises for a self-evaluation.

The courseware is composed of eight LOs each of these has been implemented as a Sharable Content Object that illustrate respectively the following fundamental basic concepts:

LO 1. Introduction

LO 2. Relational model

LO 3. Relational Languages

LO 4. Normalization

LO 5. Entity/Relationship Model

LO 6. Logical Design

LO 7. Physical Design

LO 8. SQL

Each learning object represents a single didactic unit, that students can consult through customized learning paths. The system only suggests the default sequential path for a better comprehension of the course contents.

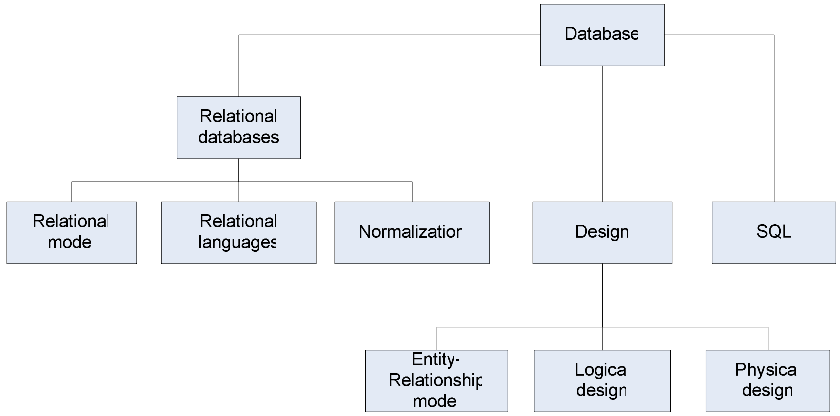


Figure 2. Database's conceptual map.

Moreover, the learning path of each student can be controlled and examined by the teacher, and the stored results of tests and exercises carefully evaluated to improve the course performances and to directly interact with single students. In fact, the system has the capability to store the student's learning path (*viz.*, visited units and result of exercises and tests), so that it is possible to recall the learning process starting from the breakpoint. It uses a database of general information about students, linked with information about the learning progress of each student (tracks on already read didactic units, exercise results, and so on …) so as to permit a self-evaluation of the obtained learning progress.

Each learning object is equipped with exercises. In fact, at the end of the theoretical contents, the student is allowed to compile a multiple-choice test for a self-evaluation. For each test, there are reported: progressive number, question, alternative answers, student selected answer, "correct/ wrong" answer flag, feedback to the student, and evaluation mark given to the student answer.

On the last row, the total evaluation mark of the entire student test session is reported. When a wrong answer is given by the student, an "help" button furnishes him with supplementary didactical information through the link to the corresponding lesson. As an example, the work flow related to the fruition of the learning object concerning the relational languages is shown in Figure 3.

B. Laboratory Lessons

Lessons held in didactic laboratories regard the study and the utilization of three of the most utilized and popular DBMSs: (a) MySQL, (b) DB2, and (c) Access. Each of them covers specific training needs and provides different skills, since they differ considerably as concerns the architecture, features, and level of difficulty.

According to our teaching experience, MySQL is the best tool to be used for academic lessons. In fact, it presents the following advantages:

- *It is open-source*. Students can easily obtain the most recent release and install it on their own Windows or Linux/Unix machines. Moreover, the chance to access the C source code makes this tool suitable for Programming courses, where students can learn to write and/or to modify C programs.

- *It is standard SQL compliant*. Databases courses require the study of the SQL language and this tool is characterized by a strong compatibility with the standard statements of this language. In addition, it gets both a graphical user interface and a command line prompt, that force students to know the most important statements by heart.

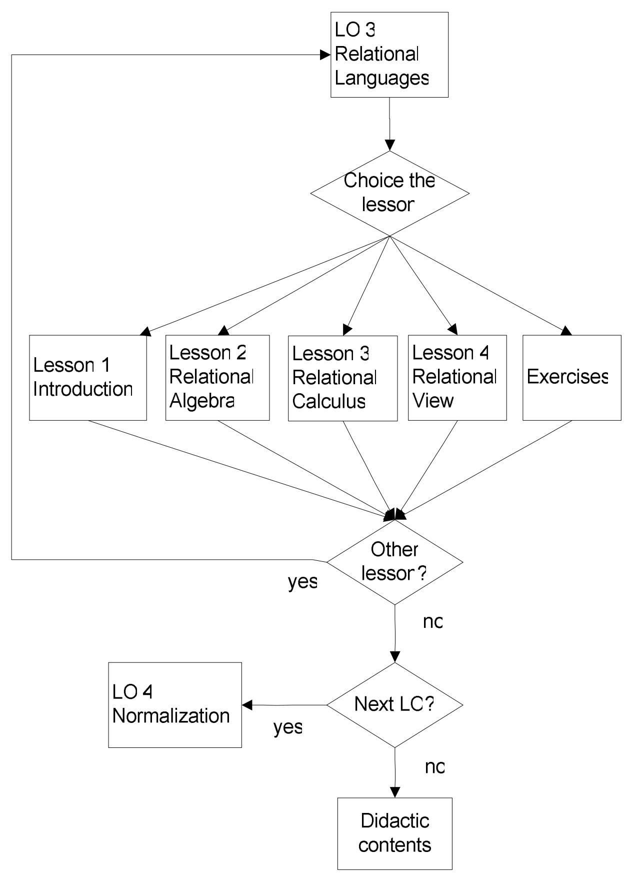


Figure 3. Lesson fruition.

- *It guarantees a stable interaction with development environments*. Nowadays, many Software Engineering courses impose the design and the implementation of a didactic project. More and more students choose to develop their own projects using web-oriented languages, like PHP or Java, and to deploy the developed applications via a web server like Apache. Indeed, MySQL is the best software to guarantee a perfect integration with these systems.

On the other hand, DB2 is a powerful DBMS, mainly used in the business company context, for Windows and Linux/Unix machines. Its advantages are:

- *It has an XML engine.* DB2 is able to manage both relational databases and XML files via a native engine. Moreover, it allows users to write SQL queries to access to both data stored in relational databases and XML files at the same time. This is coherent with the business companies workflows, that always need to integrate data coming from heterogeneous data sources.
- *It supports the SOA architecture.* Web services help business companies to deliver new products and services quickly and directly in safety, since it is based on HTTP. In addition, it allows to reduce the costs related to the delivering.
- *It simplifies the system management.* DB2 provides a set of advanced graphical tools for a centralized management of the system, the objects created by the user, and the tasks. For example, it is possible to schedule a database backup task at night and receive an email about the success of the task.

When the training requirements do not need high level of difficulty and do not intend to provide specialized skills, a good choice falls on Access, that is a widely used tool for Windows machines. It is the most used software for ECDL courses and it is part of the curricula of the Informatics teaching in the humanism-oriented degree courses for the following reasons:

- *It is easy-to-manage*. Access allows users to deploy their databases according to a rapid development time. In addition, it is very cheap to implement and to maintain databases during the time.
- *It is highly usable*. Its user interface is very friendly and consistent with the other Windows-based applications; in this way, all the users can reuse their knowledge in order to learn to perform new tasks.
- *It is easy-to-learn*. Access does not require a deep knowledge of the databases theory and it is provided with a good documentation.

However, in spite of technology and architecture, all the DBMSs have the same features in common. For the design of the learning objects of the laboratory lesson, we have defined a conceptual map that shows the main topics related to a DBMS, according to a tree structure that highlights the hierarchical relationships (see, Figure 4). In this way, we focused the attention on the main concepts, rather than the architecture of the systems.

IV. Statistical Analyses of the Content Fruition

The Databases course of the Computer Science degree course of the University of Bari is attempted by the students enrolled to the second year. This course is held from September to December, and the period suggested for the examinations is January to February.Moreover, the number of the students regularly attempting the second year of each Computer Science degree course is 70.

We solicited all of them to perform the registration at the e-learning platform and the enrolment into the Databases courseware, in order to verify whether an improvement of the didactics quality can be obtained.

Thanks to the ATutor's features, that include an useful tool that allows teachers to trace the activities of the students registered into the system, we are able to execute statistical analyses about the student's feedback.

As we expected, not all students have carried out the complete steps to start an e-learning process, since the participation to the Databases courseware is not mandatory and it represents a support tool for the face-to-face lessons.

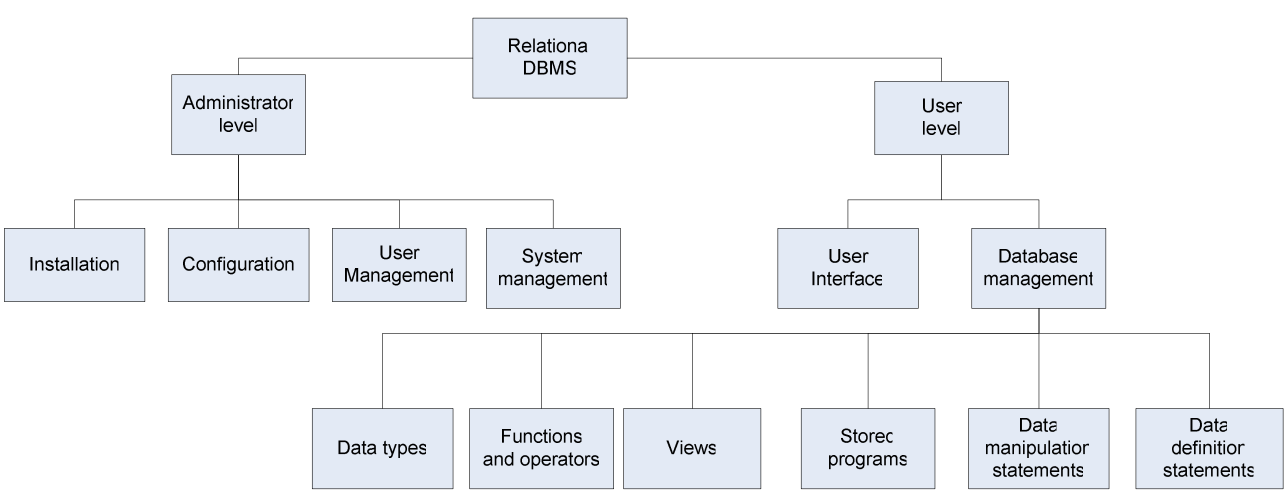


Figure 4. DBMS conceptual map.

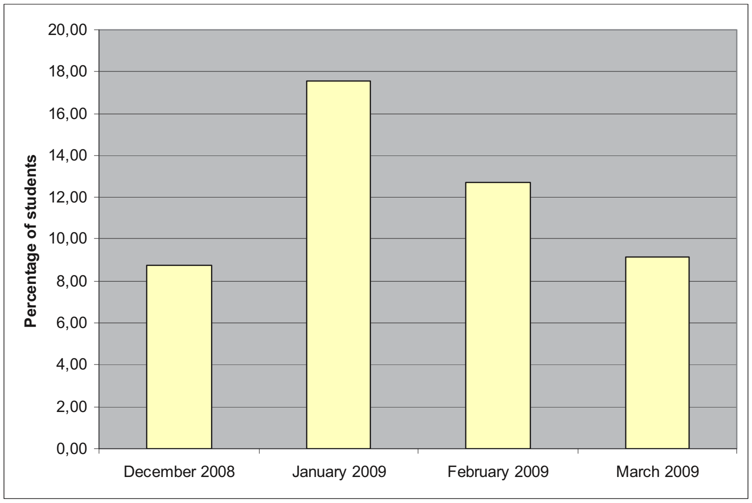


Figure 5. Students' daily frequency.

In fact, the percentage of students that have enrolled to the courseware is 40%. Further analyses have shown that most students have consulted the on-line didactics contents with a view to the examinations in January and February, and not during the period of the traditional fact-to-face lessons. In January, almost 18% of the students consulted the learning objects daily (see, Figure 5). In fact, this is the month chosen by a lot of students for the final examination, as it follows shortly the end of the lessons. During the rest of the year, the number of students who intend to sustain the examination is usually very low, and, for this reason, we expect a trend to a decreasing frequency to the courseware.

As concerns the statistics about the didactics contents, we have noted an uniform access to all the learning objects, except for the home page and the introduction that are the most visited pages. This means that every student, logged at the first time, reads the preliminary concepts, then, the next time he jumps directly to the learning objects regarding the topics of interest.

V. Course Evaluation

The evaluation of the web site of the Databases course has been based on Nielsen's guidelines [20-22], which has inspired the questions we have inserted into an on-line questionnaire. The answers freely furnished by the users can help us to identify possible causes of difficulties and to improve the course usability.

The guidelines have been established after having identified the more frequent problems arising when users access the web sites. These important guidelines can be summarized into the following ten rules:

- *Use simple and natural dialogue*. The system should be usable, without help or instruction, by a user who has knowledge and experience in the application domain but no prior experience with the system.
- *Speak the users' language*. Utilize words, phrases, metaphors, and concepts familiar to the user using its perspective.
- *Give navigational feedback*. Explain who you are and what you do. Allow the user to determine her/his current position in the document structure. Make it easy to return to an initial state.
- *Be consistent*. Indicate similar concepts through identical terminology and graphics. Be adherent to uniform conventions for layout, formatting, typefaces, labeling, etc..
- *Prevent errors*. Define constraints to avoid that the system can make possibly dangerous actions without user's confirmation.
- *Recognition versus recall*. Minimize the user's short-term memory load. Do not force users to remember key information across documents.
- *Build flexible and efficient systems*. Accommodate a range of user sophistication and diverse user goal. Lay out screens so that frequently accessed information is easily found.
- *Design aesthetic and minimalist systems*. Create visually pleasing displays. Eliminate information which is irrelevant or distracting.
- *Good error messages*. Provide guidelines for error detection and reparation. Keep informed of errors or exceptions that are relevant and of interest to the user through clear, concise, and unambiguous language.
- *Help and documentation*. Task centred minimal manuals.

According to the introduced web usability guidelines, we have determined some relevant questions to submit to the users for the on-line course evaluation. Each question is associated with a problem severity rating as follows: severe (1), serious problem (2), non-critical (3), minor inconsistency (4), no problem (5), where 1 represents the most critical error and 5 represents a situation of optimal usability. Some of the questions submitted for the on-line evaluation are shown in Table I, as an example of data we are able to gather, in order to try to detect possible sources of non-usability. The results obtained from the analysis of the answers given by the students have pointed out some rigidity in the structure of the courseware. In particular, a lot of users has expressed difficulty in the use of the exercises for the self-evaluation. For this reason, the feedback from questionnaire of the usability test accomplished by students suggests future updating to the course structure and contents.

TABLE I. QUESTIONS FOR THE COURSE EVALUATION (PARTIAL).

Question	Score
Are the hyperlinks always visible and recognizable?	1 2 3 4 5
Are the hyperlinks based on consistent logical paths?	1 2 3 4 5
Is the global map always reachable?	1 2 3 4 5
Are you conscious of your location with respect to the global map?	1 2 3 4 5
Do you cancel often your last action?	1 2 3 4 5

VI. Conclusion

In this paper, we have presented our experience in designing and using learning objects for the courseware on Databases for Computer Science degree courses. The wide demand of Databases courses, according to different levels of skills, has suggested us to organize the didactic contents in modular way, in order to support the different teaching and learning needs. The courseware has been published via the ATutor e-learning platform. The features held by this tool allowed us to trace the students' activities and to produce charts containing statistical data of the didactic content fruition. Such statistics allowed us to state that the students prefer to access the digital contents just after the end of the face-to-face lessons, that is, when they have started to study for the final examination. During the rest of the year, we expect a low frequency to the courseware, due to the few students that intend to sustain the examination. Moreover, we have noted an uniform access to the digital contents, and this trend means that every student aims to increase his/her knowledge about a specific topic. Future work is intended to extend the didactic contents with further lessons and exercises, in order to increase the frequency of the students during all the months, specially during the period September to December. In this way, we hope to integrate successfully face-to-face lessons, both theoretical and laboratory lessons, and the courseware.

References

[1] J.J. Ekstrom and S.R. Renshaw, "Database curriculum issues for four-year IT programs", in *Proc. CIEC*, Tucson, Arizona, 2003, pp. 1-9.

[2] S.D. Urban and S. W. Dietrich, "Advanced Database Concepts for Undergraduates: Experience with Teaching a Second Course", *ACM SIGCSE Bulletin*, vol. 33, pp. 357-36, 20011.

[3] M.A. Robbert and C.M. Ricardo, "Trends in the Evolution of the Database Curriculum", *ACM SIGCSE Bulletin*, vol. 35, pp. 139-143, 2003.

[4] T. Shimura, M. Yoshikawa and S. Uemura, "Storage and Retrieval of XML Documents using Object-Relational Databases", in *Proc. 10th DEXA*. Lecture Notes In Computer Science, vol. 1677, Springer, 1999, pp. 206-217.

[5] J. D. Ullman and C. Zaniolo, "Deductive Databases: Achievements and Future Directions", *SIGMOD Record*, vol. 19, no. 4, pp. 75-82, Dec. 1990.

[6] K. Won, "Research directions in object-oriented database systems", in *Proc. ninth ACM SIGACT-SIGMOD-SIGART symposium on Principles of Database Systems (PODS)*, 1990, pp. 1-15.

[7] *European Computer Driving Licence*, www.ecdl.it.

[8] T. Boyle, "Design principles for authoring dynamic, reusable learning objects", *Australian Journal of Educational Technology*, vol. 19, pp. 46-58, 2003.

[9] *ATutor Learning Content Management System*, www.atutor.ca .

[10] IMS Global Learning Consortium, www.imsglobal.org.

[11] *Advanced Distributed Learning*, www.adlnet.gov/ Technologies/scorm/default.aspx.

[12] XHTML2 Working Group Home Page, www.w3.org/MarkUp.

[13] *MySQL*, www.mysql.com.

[14] *IBM DB2 Software*, www-01.ibm.com/software/data/db2.

[15] Access Home Page, office.microsoft.com/en-us/access/FX100487571033.aspx.

[16] D.H. Jonassen, "Instructional design model for well-structured and ill-structured problem-solving learning outcomes", *Educational Technology: Research and Development*, vol. 45, no. 1, pp. 65-95, 1997.

[17] *Computing Curricula*. ACM/IEEE Computer Society, 1992.

[18] R. Elmasri and S.B. Navathe, *Fundamentals of Database Systems*, fifth edition. Boston, MA: Addison Wesley, 2006.

[19] D.M. Edwards and L. Hardman, "Lost in Hyperspace: cognitive mapping and navigation in a hypertext environment", in *Hypertext: theory into practice*. Exeter, UK: Intellect Book, 1999, pp. 90-105.

[20] J. Nielsen and H. Loranger, *Prioritizing Web Usability*. Berkeley, CA: New Riders Press, ISBN 0-321-35031-6, 2006.

[21] J. Nielsen and M. Tahir, *50 Websites Deconstructed*. Indianapolis, New Riders Publishing, ISBN 0-73571-102-X, 2001.

[22] J. Nielsen, *Designing Web Usability: the Practice of Simplicity*. Indianapolis: New Riders Publishing, ISBN 1-56205-810-X, 2000.

A Study of ‘Health Promotion Course for Music Performers’ Distance-learning Course Development

Yu-Huei Su[1], Yaw-Jen Lin[2], Jer-Junn Luh[3], Heng-Shuen Chen[4]

[1]Department of Music, National HsinChu University of Education, HsinChu,Taiwan

[2]Department of Management Information System, Central Taiwan University of Science and Technology, Taiwan

[3]Department of Physical Therapy, National Taiwan University, Taiwan

[4]Department of Medical Informatics, National Taiwan University College of Medicine. Taipei,Taiwan.

chenhs@ntu.edu.tw

Abstract: *The ‘Health Promotion Course for Music Performers’ Distance-learning Course was conducted in the spring of 2009 to improve music majors’ health literacy, prevent anxious performance/performance injuries, and conceptualize correct mental/physiological care (2-hour weekly allotment/14 weeks). The distance course covers three course structures including: A. Prerequisites (muscular/skeletal structures, fitness, mental health, preventive medicine); B. Core Course (Performers’ muscular/skeletal injury prevention, performance anxiety, hearing impairment, voice health); and C. Applied Course (Evidence-based Medicine, stress relief/relaxation, Alexander Technique). Conducted at Graduate Institute of Music, National HsinChu University of Education (NHCUE Graduate Institute of Music), the instructors are music education professors and medical experts of family medicine, psychiatric medicine, rehabilitative medicine, and Otolaryngology, National Taiwan University (NTU) College of Medicine, and 50% of the students who major in music education are music teachers. The course covers lectures and discussions as PowerCam presentation transcripts in class, and uplinked to XMS learning system for online sharing. Under the Synchronous Discussion of a Distance Learning System through Web Office, physicians at NTU College of Medicine, music teachers/students, and audit music students/teachers, National Tainan University of the Arts are able to engage in inter-school discussions inside the distance education classroom.*

Keywords: music performance, health promotion, course, e-learning

INTRODUCTION

1.Most music majors deal with ‘performance and health related problems’ solitarily or turn to instructors for help

Freshmen music majors often encounter performance related physiological problems. According to survey, 60.4% of the music majors in Taiwan deal with performance injuries solitarily or turn to instructors for help; only 21.6% seek medical attention[1]; Pain or discomfort due to poor posture or over-practice and performance anxiety are most frequently experienced by performance majors at Royal Academy of Music ; however, they tend to turn to their instrument (voice) instructors for advice rather than resorting to medical help [2]. Evidently, music instructors play a vital role in ‘Music Performance and Health Promotion.’

2.Developing Health Promotion Course for Music Performers in Taiwan to meet international appeal for HPSM

The HPSM, co-sponsored by PAMA and NASM, was implemented to develop and offer an occupational health course for music majors in accordance with a proclamation at Health Promotion in Schools of Music Conference in 2004[3]. The performance and health promotion courses conducted in music departments in Taiwan are limited in number and confined to training related improvement (e.g. Alexander Technique) and practical skills instead of medical knowledge related to performance. A small number of courses on ‘Performance Medicine’ conducted by physicians focus on medical theories, but music education and performance & teaching related issues still require music teachers’ participation in discussion. In light of this, Health Promotion Course for Music Performers in Taiwan require further development to meet international appeal for HPSM

PURPOSES

1.Health Promotion Course for Music Performers is conducted in colleges/universities to foster Students' health and performance knowledge, attitude, and behavior

Performance injuries experienced by music majors are not only related to the physiological and psychological mechanisms, they often involve social, environmental, and cultural factors in addition to music itself. Health Promotion Course for Music Performers courses therefore coincide with the preventive education and intervention principles as it is more than education and information dissemination; it focuses on values, perception, and motives of music majors. Through school intervention, students will develop a sense of responsibility for injury prevention and reduce injury-reduced discomfort. The first step to reducing the occurrence and severity of injuries is for music academies to regulate factors that affect students' conducts through values, beliefs, and actions.

2.Engaging in interdisciplinary cooperation in music education, performance & teaching, and medicine to undertake course development, implementation, and evaluation

Music teachers that specialize in music performance, music education, music teaching, or conducting play vital roles in reducing performance injuries. As music teachers' social/cultural values and beliefs have great influence on students, they should take the initiative to prevent performance injuries and establish strategies to cooperate with experts in related areas. Without the participation of music teachers, the experts' efforts will simply be in vain[3]. In consideration to the inadequacy of Health Promotion Course for Music Performers courses in Taiwan, interdisciplinary cooperation in music education, performance & teaching, and medicine in this research will aid in developing the course content, implementing teaching, and evaluating results. The consolidation of the course would not have been possible without the participation of the music teachers and collaborators.

3.Interdisciplinary team teaching through distance learning and teacher-student interactions

E-learning facilitates interactive teaching and learning through high-tech media devices at any time or place. E-learning promotes learning through production, transmission, extraction, management, and exchange of contents[4]. It is classified into 'synchronized learning'(with time restrictions) and 'non-synchronized learning' (without time restrictions). Currently, Performance and health courses in different nations have adopted auxiliary e-learning such as 'Occupational Health' in University of North Texas, and 'Occupational Health: Lessons from Music' in 'Center for Biomedical Research in Music, Texas.' With the Online and Face-to-Face Instruction, 1/3 of the course is conducted in school requiring class attendance (1-hour allotment/week) and the remaining 2/3 of the course is conducted through non-synchronous distance learning. Online courses that cover e-learning targets, simulated clinical treatment, off-campus learning resource links, and interactive survey/test platforms are available 24 hours a day, 7 days a week, and 365 days a year! Meanwhile, the face-to-face instruction is designed to supplement and extend non-synchronous online learning and enhance student interactions[5]. The Music Performance and Health Course is interdisciplinary team teaching based; thus, limitations on the time and venue of class conduction must be overcome. The purpose of this research is therefore to develop synchronized and non-synchronized 'Health Promotion Course for Music Performers' Distance-learning Courses through the online learning system to facilitate interdisciplinary teaching and teacher-student interactions.

METHODS

1.Interdisciplinary Research Team Makeup

This course (2 credits; 2-hour weekly allotment/14 weeks) is conducted at Graduate Institute of Music and the course R&D team consists of experts in music education, performance & teaching, medicine, and information. The lecturers are made up of music education researchers and instructors from other music departments including vocalists and performers who teach at music departments of education universities and art universities. The collaborators in the distance-learning course currently hold job posts in NTU College of Medicine (Family Medicine, Psychiatric Medicine, Physical Therapy, and Otolaryngology). Through years of experience in distance learning, participants of NTU College of Medicine are well experienced in e-teaching aid production and distance learning. Moreover, technical support and classroom resource sharing are also available at Distance Education Classroom of NTU College of Medicine.

2.Related course contents, teaching methods, and course evaluation methods collected for approval at NTU Research Ethics Committee

The ADDIE teaching design method is adopted to develop courses and engage in requirement analyses, design, development, implementation, and evaluation. The course requirement analyses cover students' initial learning behaviors, learning targets, media resources, and class conduction costs. The course design, development, implementation, and evaluation cover the scheduled weekly progress of the course and standardized lecturer/teaching processes. The course evaluation can be obtained through: information of students' mental/physiological state, evaluation tools used to investigate students' health and living, practice, and performing behavior changes before and after class.〔6〕Since the course involves discussions on performance injury case studies, NTU Research Ethics Committee approval will be secured.

3.Selecting an e-learning system and instruction method

Since most of the lecturers are attending physicians at teaching hospitals, they are occupied with outpatient service, ward, and teaching undertakings. Getting them involved in the course R&D without affecting work remains a problem to be overcome. To cope with the need, a synchronized e-learning system, Web Office E-learning System, is selected for distance learning. This system is used by instructors/students at National Taiwan University College of Medicine and NHCUE. Furthermore, all teaching aids and lectures are prepared on PowerCam as presentation transcripts and the 100-min class comprises 60-min lecture (knowledge imparting) and 40-min discussion (questions & answers).

RESULTS

1.Level 'Health Promotion Course for Music Performers' Distance- learning Course

The Health Promotion Course for Music Performers conducted at the following academic institutions collected in this research cover: Performance Preparation(Ohio University School of Music), Health Promotion and Prevention of Injury for Musicians(University of Indianapolis), Occupational Health :Lessons from Music(University of North Texas), Health Issues for Instrumentalists (northwestern university), Keys to Healthy Music (Eastman School of Music/University of Rochester[5]；Dynamic Posture and Alignment(University of Southern Maine), The Complete Musicians (Shepherd University), Healthy Musicianship(Michigan State University[8]；Art of Teaching Music: Wellness Practices for Musicians(George Mason University), Professional Skills(Royal College of Music), Health Courses and Training Program in Music Physiology(Hannover University of Music and Drama) [9].

In light of the above, the 'Health Promotion Course for Music Performers' Distance-learning Course developed by the research team covers three course structures including: prerequisites, core course, and applied course.

The prerequisites cover: basic mental/physiological health, the core course introduces mental/physiological issues related to performance and health, and the applied course covers the practical skills that promote performance and health. Moreover, graduate students are educated on ways to determine the evidence level of medical references and evidence-based medicine in performance is scheduled to enhance the graduate students' research competence. The prerequisites cover: muscular/skeletal structures, fitness, mental health, preventive medicine; the core course covers: performers' muscular/skeletal injury prevention, performance anxiety, hearing impairment, voice health; and the applied course covers: Evidence-based Medicine, stress relief/relaxation, Alexander Technique.

2.'Health Promotion Course for Music Performers' prepared as Presentation Transcripts on PowerCam

PowerPoint is widely used in teaching for its user-friendly and good results. Presentation contents are often the essence of a piece of knowledge, but can only be saved as a PPT file at the end of a presentation instead of the lecture itself. Through PowerCam (presentation/screen video recording software) on PowerPoint, the lecture process is produced into e-learning teaching aids (including the lecturer's image, voice, lecturing frames, mouse cursor, etc.), the course content is copied through PowerCam, and uplinked to XMS Learning System for internet sharing.

3.Synchronized Distance-learning for Physicians, Music Teachers, and Students through Web Office E-Learning System

Web Office, a multi-User / multi-Point/ multimedia conference service system developed by Far Ancient Tech., is an effective synchronized online face-to-faces system adopted in the course. Physicians at Distance Education Classroom of NTU College of Medicine and Distance Education Classroom of NHCUE engage in teacher-student discussion (Figure1). Audit teachers/students of National Tainan University of the Arts are also able to access the ongoing class online and interact with the physicians live using a microphone and video camera. Web Office features multimedia file functions such as whiteboard and shared browsing.

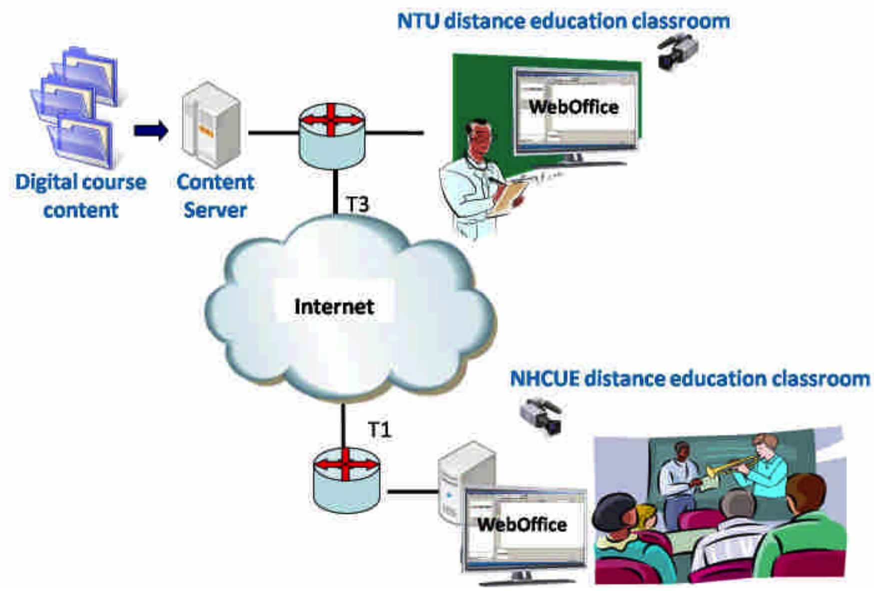


Fig. 1: Face-to-face of 'Health Promotion Course for Music Performers' Distance-learning Course

DISCUSSION

This distance-learning course assists performance related professionals in occupational health and injury prevention. The R&D courses also serve as reference for other performing arts departments and are valuable in promotion and application-based undertakings. The implementation results may be compared with those of other nations in preparation for 'HPSM' international exchanges.

The United States dominates in e-Learning applications of Health Promotion Course for Music Performers over other nations. Our information education-based nation integrates e-learning and course teaching to enhance the performance and health literacy and e-learning capabilities of music majors and course promotions in the future. Currently, Japan is still the more active participant of performance medicine related research among other Asian nations, but our nation's academic visibility can be enhanced through sustainable development as a team.

The e-learning platform in this research facilitates interdisciplinary team teaching, Health Promotion Course for Music Performers Distance-learning Course completion, participants' online course development, and three-way interactions (physicians, teachers, and students). The course focuses on participants' interdisciplinary knowledge acquisition. More importantly, the professionalism interdisciplinary integration can be strengthened and the academic horizons can be expanded.

REFERENCES

[1] Peng Yi-wen, Sun Pei-chang, Chang Hua-an, Hsu Dao-chang (1994), Preliminary Report of Professional Injuries of Music Majors, Journal of Rehabilitation Medicine Association, 22(1),105-109。

[2] Williamon, A. & Thompson, S.(2006). Awareness and incidence of health problems among conservatoire students. Psychology of Music, 34(4), 411-430.

[3] Chesky K.S, Dawson W.J.,& Manchester R.A.(2006). Health promotion in schools of music: initial recommendations for schools of music. Medical Problems of Performing Artists, 21(3),142-144.

[4] Manpower Planning Division, Council for Economic Planning and Development (2001). The Present and Future of E-Learning, December 25th, 2001. Press Release.

[5] Manchester, R. A.(2007a).Health promotion courses for music students: part Ⅰ. Medical Problems of Performing Artists, 22(1),26-29.

[6] Rebecca Barton, Judy R. Feinberg (2008). Effectiveness of an Educational Program in Health Promotion and Injury Prevention for Freshman Music Majors. Medical Problems of Performing Artists,23(2),47-53.

[7] Manchester R.A.(2006). Toward better prevention of injuries among performing artists [editorial]. Medical Problems of Performing Artists, 21(1),1-2.

[8] Manchester, R. A.(2007b).Health promotion courses for music students: part Ⅱ. Medical Problems of Performing Artists, 22(2), 80-81.

[9] Manchester, R. A.(2007c).Health promotion courses for music students: part Ⅲ. Medical Problems of Performing Artists, 22(3),116-119.

Understanding Art Exhibitions: From Audioguides To Multimedia Companions

Giuseppe Barbieri[1], Augusto Celentano[2], Renzo Orsini[2] and Fabio Pittarello[2]
[1]Dipartimento di Storia dell'arte e conservazione dei beni artistici, Università Ca' Foscari, Venezia, Italia
figaro@unive.it
[2]Dipartimento di Informatica, Università Ca' Foscari, Venezia, Italia
{auce,orsini,pitt}@dsi.unive.it

Abstract

This paper elaborates on mobile devices for assisting the visitors of an art exhibition to understand the exhibition and its content. A model for a new class of multimedia guide is described, instantiated into a prototype for an exhibition on sacred Ethiopian art in Venice. The model offers several levels of use to different classes of users. A prototype guide has been evaluated through questionnaires and traces of users' exploration, automatically extracted from the guide records.

1 Introduction

Understanding art requires education; even the supposed evidence of meaning of Classic art and Renaissance painting (to cite only two easy examples) requires a museum visitor to be knowledgeable about the author and to understand elements of composition, history, methodology of analysis, detail interpretation, which are at the core of courses on art history and art criticism.

Modern and contemporary art, loosing the aesthetic resemblance between the artwork appearance and the subject, and multiplying the materials and techniques, makes more evident the need for organized and methodic knowledge about the artist, his/her historical, political and cultural context, the used techniques and their role in the society, the overall context in which the artwork has been conceived, made and exposed to the public.

Generally speaking, any form of art presents so many facets that its fruition cannot rely on the visitor immediate perception only, but requires (at different degrees) explanations and comments to be understood and enjoyed. In museums a basic level of knowledge is provided by two common devices: the large panels introducing the artworks of a room, of an author, or of a section, and the (small) labels aside each artwork with basic data such as title and date and, sometimes, a short comment.

Labels have been introduced by Tommaso Puccini, the superintendent of the Uffizi Galleries in Florence, who lived between 1749 and 1811; around 1780, he decided to place a label aside each artwork with the name of the artist, the subject, the execution date and the technique used. This practice ended an era when fruition of art was a pure aesthetic process, in which the visitor could recognize the masterpieces but was otherwise attracted by other issues such as the vastity and variety of a collection, the personal ability to discover details, and so on. The labels gave the visitor the ability to *know* rather then to *experience*, replacing knowledge for pleasure: who is the author, what is the content (often based on allegories or historical events not immediately evident), what inspired the author (e.g., a replica of another artwork). The personal experience in appreciating an artwork, often in company of the collectionist hence reserved to few people, was replaced by a collective, accessible knowledge of objective information.

The history of art and of art criticism has been greatly influenced by this little revolution, but a discussion on such issues would lead out of the scope of this paper, whose focus is on the possibility (the *need*) of regaining (part of) the ancient esprit that led visitors to art collections, and that current ITC can help to establish.

It is generally acknowledged that today's attitude in visiting art exhibitions (mainly with masterpieces) is neither to know nor to experience, but rather to *recognize* and to *recall*; hence, information provided to users must support knowledge elicitation, which should drive the artwork examination process after recognition has been made.

In the remainder of this paper we shall discuss the main issues about the design and the evaluation of a new type of multimedia guide devoted not only to present facts about the exhibition and the artworks, but mainly to help the visitor to understand what he/she is visiting, integrating factual information with the proper cultural context. The audiovisual channels used for communicating information might be used also for addressing the emotional side of the visit, giving the user the sense of a more complete experience.

A prototype of such a guide, based on the Apple iPod touch, has been experimented for an exhibition about religious Ethiopian art, "Nigra Sum Sed Formosa - Sacred and Beauty in the Christian Ethiopia", held at Ca' Foscari in Venice, Italy from March 13 to May 10, 2009 (http://www.nigrasum.org/).

The guide has been conceived in the framework of a research project about new interactive systems for rich art fruition involving the Department of Computer Science and the Department of History of Arts and Cultural Heritage Preservation at Ca' Foscari University. The goal of the project is to design multimedia guides that do not address visitors only with simple, albeit correct, comments on the artworks on display, but rather provide ample knowledge about the exhibition themes.

Such a goal requires a balance between the simple linearity of a guided tour aimed at explanation and a rich hypermedia structure aimed at involving the user emotionally. Indeed, it is a goal common to all museum and exhibit guides, but the ways it is claimed or pursued are varying and, at our knowledge, often limited by the existing habits and practices about audio guides.

2 New directions for art guides

From a humanistic perspective a good guide should give visitors not only information, but also, and mostly, experience coming from the interaction with the artworks. Multimediality, the technological key to unreveal such experience, cannot be limited, as often done, to provide impressive views on information fragments; it must be taken as a way to allow the visitor to enter the artwork meaning from a personal emotional point of view. Using interactive multimedia portable devices delivering information in the shape of engaging audiovisual presentations is a starting point and not a goal [3, 5].

From a technical perspective designing such a guide faces several challenges, the most notable being the almost unreachable ease of use of the widespread and simple, albeit limited, keypad-based audioguide. Any enrichment leading to more choices than listening to an audio comment raises problems of use, dividing the users according to their skill with portable devices. The operations of an audioguide are explained and learned in seconds, while any PDA class device requires minutes, a time span unbearable in large museums.

Indeed, PDA class devices are being introduced in museums visited by a large public [2]. An interesting case is the Ship Museum in Barcelona (http://www.mmb.cat) that provides visitors with a touch smartphone of the Windows Mobile class. The PDA functions are constrained to the selection of a presentation typing on a numeric keypad the number marking the proper museum section. Each presentation is an audio comment, accompanied by additional content in the shape of images and short texts. As an additional bonus, the visitors can bookmark relevant images and texts, that will be sent to their e-mail address at visit's end. The bookmarked information is sent as a simple list of attachments, without any structure or reference to the museum context.

The Apple iPod touch is changing the landscape of interactive guides, due to the excellent quality of audio and video playback, to the interaction style and to a fashionable appeal of the device itself. IPod touch based guides are available at several museums, among which the New York MoMA and the Tate Liverpool Gallery are notable examples. Indeed, in both cases the iPod is used as an unstructured mediatheque. At MoMA it is promoted as the *MoMA Audioguide* (http://www.moma.org/visit/plan/atthemuseum/momaaudio) delivering audio-only comments with limited interaction capabilities. At Tate Liverpool an exhibition on Klimt masterpieces was supported by a catalog of audiovisual presentations with very simple visual content (http://www.tate.org.uk/liverpool/exhibitions/gustavklimt/tour.shtm). None of the two guides uses at full extent the video capabilities of the iPod to enrich the vision of the artworks, nor they use gestual interaction out of list selection.

Also the Öberes Belvedere Gallery in Vienna (http://www.belvedere.at) delivers guides on iPod touch devices, which "simulate" a traditional keypad audioguide. The user interacts with a virtual keypad displayed on the iPod screen selecting artworks by number, and listens to audio comments with very limited playback control: there are no advance/back controls, the volums is controlled with the hardware buttons on the iPod side, and the only allowed operations are pause/resume and return to the keypad for a new selection. Such a "downsizing" of the iPod capabilities turns a sophisticated multimedia device into a very basic audio player.

3 The art guide as a visitor's companion

The goal of the guide we have designed is to help the visitor to understand the value (artistic, historical, religious) of the artworks and of the exhibition as a whole while being on site. Information access modalities support different styles to visit the exhibition, including non-linear visits that—according to [4, 7]—represent the behaviour of a significant part of visitors. Figure 1shows the organization of the guide content as it appears through the user interface. The user can access the content by selecting a section of the exhibition, a specific artwork from a catalog, a room on a map, or a keyword from a tag cloud.

Access by section give users information about six exhibition themes: Icons, Crosses, Devotional objects, Religious architecture, Testimonials, and Drawings. The first three sections deal primarily with artworks of a same type.

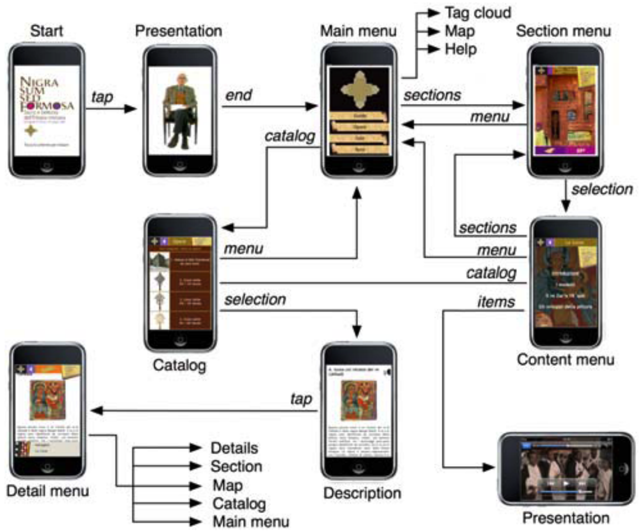


Figure 1. Organization of the guide content

The other three sections include artworks of different types, homogeneous as to the context to which they refer. Accessing the guide content by section exploits the potential of the device, with multimedia files such as movie fragments, audio recordings, images, interviews, traditional religious songs, and so on.

The artworks are accessed by several catalogs: a general catalog collecting all the artworks, and partial catalogs for the artworks of a section or located in a room. Each artwork is presented with an audio comment (much as in a conventional audioguide), with one or more images that can be zoomed in to discover details; in some cases a longer audiovisual presentation shows more information through animations and slide shows. A set of maps allows visitors to access the artworks contained in each room.

User may also access information by selecting keywords from a *tag cloud*, an information structure borrowed from the so-called Web 2.0 and often used to present the result of the collaborative tagging of users [6]. In this guide the tag cloud is used to group multimedia presentations by evocative words that are orthogonal to the exhibition sections.

In order to be used by a wide range of users with different skills and attitudes towards both the Ethiopian culture and the use of personal devices, the guide has been designed for being both engaging and usable, fitting the visitors background without being trivial. The graphic appearance has an important evocative role in suggesting atmospheres and themes. Real images from the artworks have been used to identify the different contents, instead of symbolically styled icons. Navigation is kept at a minimum level of complexity, avoiding deep hierarchies and paths.

Interaction has been designed for being accessible—in its basic functionalities—also by unexperienced users. As a consequence, the standard gesture of tapping on the screen (the equivalent of the click operation on a desktop interface) has been extensively used for selecting items and accessing information in menus, indexes and pages. For example, tapping on the image of an artwork page reveals a popup menu leading to supplemental audiovisual content. Browsing through artworks is accomplished by the standard iPod gesture of flicking the pages. More complex gestures—such as the enlarging gesture operated with two fingers—have been used only for zooming into images. The evaluation of the guide use, discussed in Section 4, shows that such gesture, although intuitive in principle, requires a learning phase the prevents many visitors to use it.

Table 1. Questionnaire evaluation - part 1

(a) Guide chapters

	Catalog	Rooms	Sections	Keywords
Used by	85%	81%	69%	66%
Score (1–5)	4.1	4.1	4.0	4.1

(b) Interface functions

	Touch	Scroll	Flick	Enlarge
Used by	86%	85%	80%	53%
Score (1–5)	4.2	4.4	4.2	4.1

(c) Content type

	Audio comments	Audio-video	Artwork pages
Evaluated by	81%	78%	82%
Score (1–5)	3.9	4.1	4.1

4 Evaluation of the guide use

We released three different versions of the guide. The second version, issued shortly after the exhibition opening, was improved mainly with the content and with a more evident identification of the guide chapters. The third version of the guide included a context-dependent audio-visual helpand further improvements in the navigation structure. In order to evaluate the effectiveness of the guide we adopted two different tools: a questionnaire and an automatic tracking system of user gestures. All the data collected are related to the second and to the third version of the guide.

4.1 Questionnaire

The questionnaire was submitted to the visitors except in hours of great affluence. Table 1 presents a synthesis of the analysis on a sample of 176 questionnaires. A general appreciation of the guide is evident. However, it is perceivable from Table 1(a) how the habit of using sequential audio guides influences the visitors that, even in presence of a more articulated device, privilege direct catalog access to introductions and theme explanations. It is also evident from Table 1(b) that the more advanced user functions, such as multitouch gestures to zoom into images, present some difficulties, and a consistent part of the users didn't even realize that such function was supported. The results suggest that multitouch gestures, associated to power and ease of use in advertising this class of devices, is still not perceived as a natural interaction and may require a learning phase hard to be satisfied in the short time of an exhibition visit.

We may expect that, with the rapid spreading of touch-based devices and the standardization of gesture-based interaction, users will be aware of these functions without an explicit help. The lesson learned for the immediate time is that we need to support the users with explicit information about the functions available.

Table 1(c) shows a satisfactory appreciation level for the different types of content provided by the guide, with a slight preference for the audiovisual content of the introductory sections and the artwork descriptions, compared to the audio-only comments associated to the artworks.

Additional information comes from questions related to the ease of use of the guide and to the availability of help tools, summarized in Table 2 for versions 2 and 3 of the guide. Only a minor part of users where often in trouble when accessing information, as shown in Table 2(a). Most of them were only sometimes bothered by access problem. A minor, but significant part of users (i.e. 19%), declared that they never had problems with the second version of the guide. The improvements to the third version of the guide (the availability of a *back* button and the integrated help) increased this percentage to 28%.

Table 2(b) shows data about the use of the help: a leaflet in version 2 and an integrated video presentation in version 3. As expected, the video help was more appreciated, and was considered useful by the 68% of the users and a must from an additional 10%. These percentages represents a considerable improvement if compared to the appreciation of the paper leaflet, that can be explained in particular by the contextual access from any section of the guide and by the use of a video presentation showing the real use of the device. The printed help had also the problem that in some cases, due to visitor affluence, leaflets were not available; a poster displayed near to the iPod counter supplied the same information, but the high percentage of "no answer" in the first row of Table 2(b) is a signal of its inadequacy.

Table 2(c) displays the user responses to difficulties about navigation in the guide content. For the second version of the guide returning to the starting screen was the preferred solution, followed by asking assistance to human guides available in each room. For the third version of the guide these solutions were still the most used, but with smaller percentages. It is interesting to note that for this version of the guide a minor but significant part of the users relied on the video help for recovering from troubles.

Table 2. Questionnaire evaluation - part 2

(a) Ease of use

	No answer	No difficulty	Some difficulty	A lot of difficulties
V.2	7%	19%	63%	11%
V.3	4%	28%	54%	14%

(b) Help*

	No answer	Useless	Useful	Necessary
V.2	49%	10%	35%	5%
V.3	14%	8%	68%	10%

* A leaflet for version 2 and an integrated video for version 3

(c) Management of difficulties

	No answer	Home	Help	Go on	Ask	Other
V.2	23%	50%	n/a	14%	25%	2%
V.3	32%	38%	14%	14%	18%	3%

(d) Request for Additional Support Functionalities

	No answer	Help	Back	Home	Other
V.2	27%	15%	42%	27%	1%
V.3	47%	4%	22%	21%	7%

Finally, users were asked about their preferences for additional functionalities to improve the ease of use (Table 2(d)). A significant percentage of users of the second version of the guide (42%) requested the availability of a back button to retrieve their previous steps. Such functionality was implemented in the third version. The other answers show that, for both versions of the guide, a part of the users requested additional functionalities that were already available. This problem might come from the graphical interface style: rather than using standard traditional symbols for the *home* and *back* buttons, we used symbols and colors consistent with the graphic style of the exhibition installation. Some users might have considered these symbols as decorative elements rather than functional icons, missing their presence as navigation aids. A final encouraging result shown in Table 2(d) is that an increasing part of the users of the third version of the guide (47% vs. 27%) expressed no need for additional support functionalities.

4.2 User behaviour analysis

User behaviour analysis is based on the automatic recording of user activity. We have collected more than 100.000 records, each corresponding to a user gesture on the screen or to a file access, counting for more than 900 different visits in a time span of about nine weeks. We plan to use data mining techniques to analyze a so huge collection and to reveal recurrent user patterns; a preliminary synthetic analysis of the traces has already given us precious knowledge about the real use of the guide.

A first result is that approximatively 15% of the visitors

Table 3. In and out paths from the guide sections

	Home	Catalog	Sections	Maps	Tags	Seclist	Maplist	Video	Artwork
Home	-	1.669	1.254	3.035	928	-	-	-	-
Catalog	509	-	-	-	-	-	-	-	1.634
Sections	1.187	-	1.003	-	-	128	-	1.524	-
Maps	744	-	-	4.333	-	-	3.678	-	-
Tags	932	-	-	-	1.219	-	-	1.290	-
Seclist	94	-	-	-		-	-	-	57
Maplist	1.297	-	-	-		-	-	-	3.088
Video	-	-	1.508	-	1.275	-	-	-	-
Artwork	2.141	441	-	1.416	-	22	716	-	14.786

who took the guide has used it for less than ten minutes, while 12% of the visitors uses it for more than two hours. We do not have data on the average time of visit, but the exhibition occupies only nine rooms, and two hours are an ample time. Overall, the average time of guide use is around 52 minutes, distributed according to the Zipf law.

Considering the visit paths, we divided the guide pages in homogenous sets, and studied the in/out-paths from a set to another. The result for more than 600 visits from the second version of the guide is shown in Table 3.

The sets are: *Home*, the initial page; *Catalog*, the list of the artworks; *Sections*, the pages related to the exhibition themes, with links to the audiovisual content; *Maps*, the pages related to the exhibition rooms; *Tags*, the pages accessed by keywords from the tag cloud; *Seclist*, the list of artworks related to each section; *Maplist*, the list of the artworks present in each room; *Video*, any audiovisual content; *Audio*, the audio comments of the artworks, *Artwork*, the pages containing the description of the artworks. A cell at row i and column j contains the number of times a page in set j has been reached from a page in set i. For example, artworks pages (column *Artwork* are reached primarily from lists associated to the room maps (3088 times), then from the general artwork list (*Catalog*, 1634 times), and very seldom from the lists associated to the themes pages (57 times); a large number of times they are reached from other artwork pages, confirming that the users have used frequently the flicking gesture to advance to the next artwork.

From Table 3 a few important facts come immediately into evidence: (1) from the home page the preferred next page is the map, then the catalog, then the theme index and the tag index; (2) the audiovisual presentations have been reached almost equally often from the sections' pages as well as from the tags' ones; (3) the partial catalogs of each section have been scarcely used, while those associated to the exhibition rooms have been heavily used; (4) from the artworks pages, besides browsing sequentially the artworks, the users have returned often to the home page, to the room map and, sometimes, to the artwork catalog. According to data not shown in Table 3, audio comments associated to artwork pages have been played in the 30% of the cases.

From these facts a few conclusions can be preliminarily drawn: (1) the visitors have preferred the "utility" part of the guide, i.e. the maps and the room catalogs; (2) the visitors have exploited the audiovisual content of the guide in a significant way, with a preference for the video presentation with respect to the audio only comments; (3) in general the interface should be simplified, since the users have sometimes followed non efficient paths, like returning to the home page to continue the visit: this is confirmed also by the questionnaires.

5 A family of guide models

Tuning the guide to the visitor abilities with personal devices emerged as a requirement during the exhibition after the examination of the early questionnaires. While they showed that visitors were generally satisfied with the guide, some visitors returned the device a few minutes after picking it up, saying it was too difficult to use. Such visitors where mostly elder persons not used to personal devices. In a few cases they expected a more conventional guide, with some kind of "next" button to advance in the exhibition, making up a simple guided tour.

We cannot detail here, for space constraints, the technical solution adopted for the management of the guide content; the reader is referred to a previous paper [1] for a deeper presentation. We overview here the overall information organization to understand how the architecture we have developed can be used to build a *family* of art guides with the same basic content units but with different cognitive paths between them and different presentation structures and styles, designed to accomplish the need of users with different expectations, different abilities in using personal devices, and different experiences with art guides.

Content management is based on a database organizing the artwork data in collections. Each artwork belongs to many collections: the general catalog, the set of artworks contained in a room, the set of artworks of a guide section, and so on. The database manages also the references to multimedia files which are the ultimate components of the guide content; three types of multimedia materialexist: (1) audio recordings, which are played when the user selects the audio function in an artwork page, like a conventional audio guide; (2) audiovisual presentations related to general information about a theme, accessible from any of the of six sections and from selected artwork pages; (3) interviews with the exhibition curators, introducing the main themes of the exhibition.

It is possible to design, on such material, at least four different guide models, simpler than the one we have built, assisting the users in different ways both with respect to the operation complexity and to the information richness.

Anyway, due to the limited exhibition duration, they were not implemented at full extent but only at demonstration level in the iPod devices. The four guide models are: (1) an audio catalog; (2) an audiovisual catalog; (3) a guided tour; (4) a guide personalized on the visitor type.

Audio catalog. The set of audio presentations which are accessed through the catalog section of the guide can be stored in the iPod *Music* section, collected into a playlist with tunes properly numbered to match the artwork identifiers. The guide works as a traditional audioguide except that, instead of typing the artwork number on a keypad, the comment is selected from a scrolling list. More than one catalog can be accessed using several albums and playlists, each collecting a subset of the audio comments according to a thematic (section) or spatial (room) criterium.

Audiovisual catalog. The extension from audio-only comments to audiovisual comments turns the audioguide into a visual guide; however, the guide organization is the same as the previous case. The presence of video allows the guide designer to present introductory material, to show and compare artworks, to address the visitor with detailed visual analysis not available in an audio-only guide. In both guide models, the previous one and this one, the visitor gestures are limited to list scrolling and unitouch selection, avoiding troubles in commanding the guide functions.

Guided tour. The guided tour model is an extension of the multiple collections model that can be implemented using different playlists. The sequential activation of the tour step is built into the iPod control system for audio and video, which allows the user to go to the previous or next item of a playlist at play end, or to skip to next or previous step interrupting the playback. The tour path is represented as a list, that can be accessed at any intermediate point since the basic iPod controls allow the user to choose the item to start with.

Personalized guide. This model opens a wide fan of possibilities due to the many ways in which content, presentation and sequence of information can be combined to fulfill the needs of different categories of users. In the context of this paper we simply note that personalization can rely on the basic mechanism of aggregation of atomic multimedia presentations into sets, called catalogs, playlist, chains, tours, according to the semantic context in which they are used. The sets are accessible through two basic mechanisms: direct selection and previous/next selection, both implemented as basic controls in the iPod touch software.

6 Conclusion

Designing an art guide is an educational activity that must consider the relationship between the visitors and the artworks in designing what content must be delivered and what freedom the users have in accessing it. The relationships are much more complex than those suggested by the linear paper guides and audioguides usually available at museum entrances. Yet, the richness of the personal interactive multimedia technology risks to divert the user attention from content to presentation. Different art contexts and themes require different styles for presenting an art collection (not to speak about *explaining*, which is a debated subject expecially in contemporary art exhibitions) helping the visitors to access it with satisfaction.

The guide structure and the run-time environment we have implemented can be adapted to other contents with limited changes, making this project a first step towards the building of a generic content management system for portable multimedia guides. Indeed, we are currently developing an adapted version of the guide for a part of the exhibition ”Topological gardens”, organized by the Philadelphia Museum as the USA official participation to the Venice *Biennale* of 2009 and dedicated to Bruce Nauman. The huge differences in the exhibition themes, the number of artworks and the way they are presented and commented provide a good test bed for evaluating the design of the guide and the appreciation of the visitors.

Acknowledgments

Valeria Finocchi has coordinated the content preparation. Marek Maurizio has contributed to the implementation of the content database. The project has been supported by Banca Popolare Friuladria - Crédit Agricole.

References

[1] A. Celentano, R. Orsini, F. Pittarello, G. Barbieri. Design and Evaluation of a Mobile Art Guide on iPod Touch. *Interaction Design & Architecture(s)*, n. 5&6, 2009.

[2] A. Damala, H. Kockelkorn. A Taxonomy for the Evaluation of Mobile Museum Guides. In *MobileHCI '06: Proc. of the 8th Conf. on Human-computer Interaction with Mobile Devices and Services*, Helsinki, Finland, 2006.

[3] A. Manning, G.L. Sims. The Blanton iTour - An Interactive Handheld Museum Guide Experiment. In *Museum and the Web*, Arlington, VA, USA, 2004.

[4] P. Marti, A. Rizzo, L. Petroni, G. Tozzi, M. Diligenti. Supporting Emotional Engagement in Art Settings with Adaptivity. In *Workshop on Adaptive Design of Interactive Multimedia Presentation for Mobile Users*, 1999.

[5] O. Stock, M. Zancanaro (eds.). *PEACH: Intelligent Interfaces for Museum Visits*, Springer-Verlag, Berlin, 2007.

[6] T. Vander Wal. Folksonomy. http://vanderwal.net/folksonomy.html, 2005.

[7] E. Veron, M. Levasseur. *Ethnographie de l'Exposition.* Bibliothèque Publique d'Information du Centre Georges Pompidou, Paris, France, 1983.

A Pilot Study of e-Music School of LOHAS Seniors in Taiwan

Chao-Hsiu Lee [1], Yen-Ting Chen [2], Yu-Yuan Chang [3], Yaw-Jen Lin[4], Jer-Junn Luh[5], Hsin-I Chen[6]

[1] Music Department, Tainan National University of the Arts, Taiwan.

[2] Seattle Central Community College, USA.

[3] Hulin Primary School, Taiwan

[4]Department of Management Information System, Central Taiwan University of Science and Technology, Taiwan

[5]Department of Physical Therapy, National Taiwan University, Taiwan

[6]Department of Music, National Taitung University, Taiwan.

ychendavid@gmail.com

Abstract- *The purpose of the current study is to develop an e-Music School of LOHAS Seniors-including participants' interaction, course development, and the feedbacks from learners and their family. A group of healthy seniors in senior care centers are invited as research subjects. The development of musical learning activities is based on the Health Classification of Seniors; by studying benefits contributed by the course, it is possible to gather information for future development of functional courses for seniors with different physical conditions. With a multidisciplinary collaboration, Tainan National University of the Arts (TNNUA) will be responsible for the course design and teaching practice, the e-Music School website will be established by the Health e-learning network (HeN) team of Ministry of Education, and Taiwan Association of Gerontology and Geriatrics (TAGG) will perform the evaluation of health function before and after the music learning courses.*

Keywords: e-learning, music learning, seniors.

1. Introduction

Due to better life quality, better living environment, and 99 percent of population safeguarded by Health Care, that the average life span in Taiwan has been increasing ever since. According to Census Statistics on Population from Ministry of the Interior in Taiwan, 7 percent of the national population in 1993 is senior citizens, which fit under United Nations' definition for aging country. The National Health Research Institute in Taiwan had asserted that in 2020, 14 percent of the national population will be senior citizens〔1〕; with the increasing impact of Shoshika (decline of birth rate), the seniors in Taiwan may face the shortage of medical care.

Although senior care (e.g., senior care centers), senior medicine, and researches on seniors (e.g., university departments and researching institutions on senior study) have been developed, these developments tend to be focused on care for mental illnesses such as melancholia〔2〕. When it comes to the field of senior physical health and illness prevention, the development is still premature〔3〕, Also, the current dynamics of researches tend to put emphasis on seniors with illnesses, and lack the study on aiding those healthy seniors to maintain or achieve a better health condition〔4〕. Lastly, the expertise of the field isn't developed, and results in the lack of skilled personnel on the job. The previous points were what the researching team considered, aspects of senior care in dire of further development.

For seniors, losing controls of body functions has brought along multiple diseases. To find an effective way to prevent, intervene, or slow down the process of losing body functions, is an important topic to develop in response to the society; it is also a policy promoted by public health institutions in priority. The project targets at the healthy seniors, or seniors with less chronic diseases. By using the assessing tool, it tried to observe positive and negative effects in seniors' physical, mental, social aspects after half year of musical course. In addition, the project ruled out the many assessing tools designed for issues regarding senior health, and focused on senior mental/physical disease prevention, and mental/physical health improvements.

Currently, music related courses have been a part of planned activities in senior care center within the country and abroad; they are also included in the lectures in colleges and university for seniors. Most of these courses were scheduled to be short termed, and seldom did they appear as systematically developed materials. Contemporary courses center on basic appreciation and hands-on lessons, and were not designed with regularity, long termed progress, and depth. Besides that, these courses did not follow the Health Classification of Seniors; therefore, they couldn't aim at the different needs of the seniors. Furthermore, the courses were only upon the level of practical courses, and could not provide services of e-learning. For the family members, they could not be informed about the learner's condition in class; this reduces the effectiveness of course promotion. In response to the previously mentioned issues, the researching team intends to target at the healthy seniors in senior care centers, and create adequate musical learning activities in accord with Health Classification of Seniors to build up the practical courses of Music School of Seniors. At the same time, construction of systematic E-Music School of LOHAS seniors will be taken place, providing interactions within the community of the learner and the teachers, course information and materials, and the feedbacks from the learners and their family. This project carries out sets of evaluations of results for practical courses and e-learning, and studies the positive effect of the course on senior health. In the long run, the hope is to provide constructive suggestions for future development of functional courses for seniors with different physical conditions.

2. Method

2.1 Framework

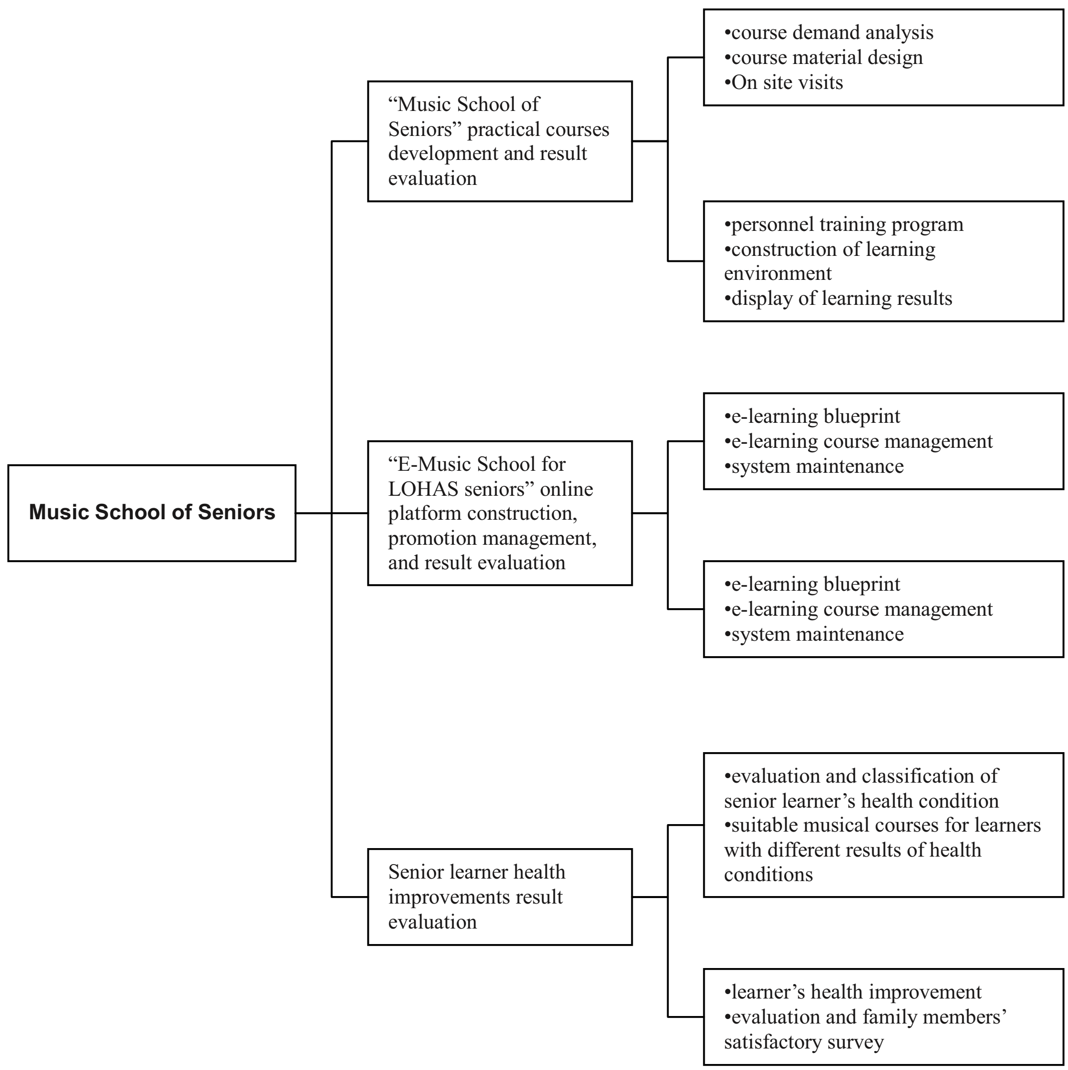


Fig. 1 System Framework of e-Music School of LOHAS Seniors

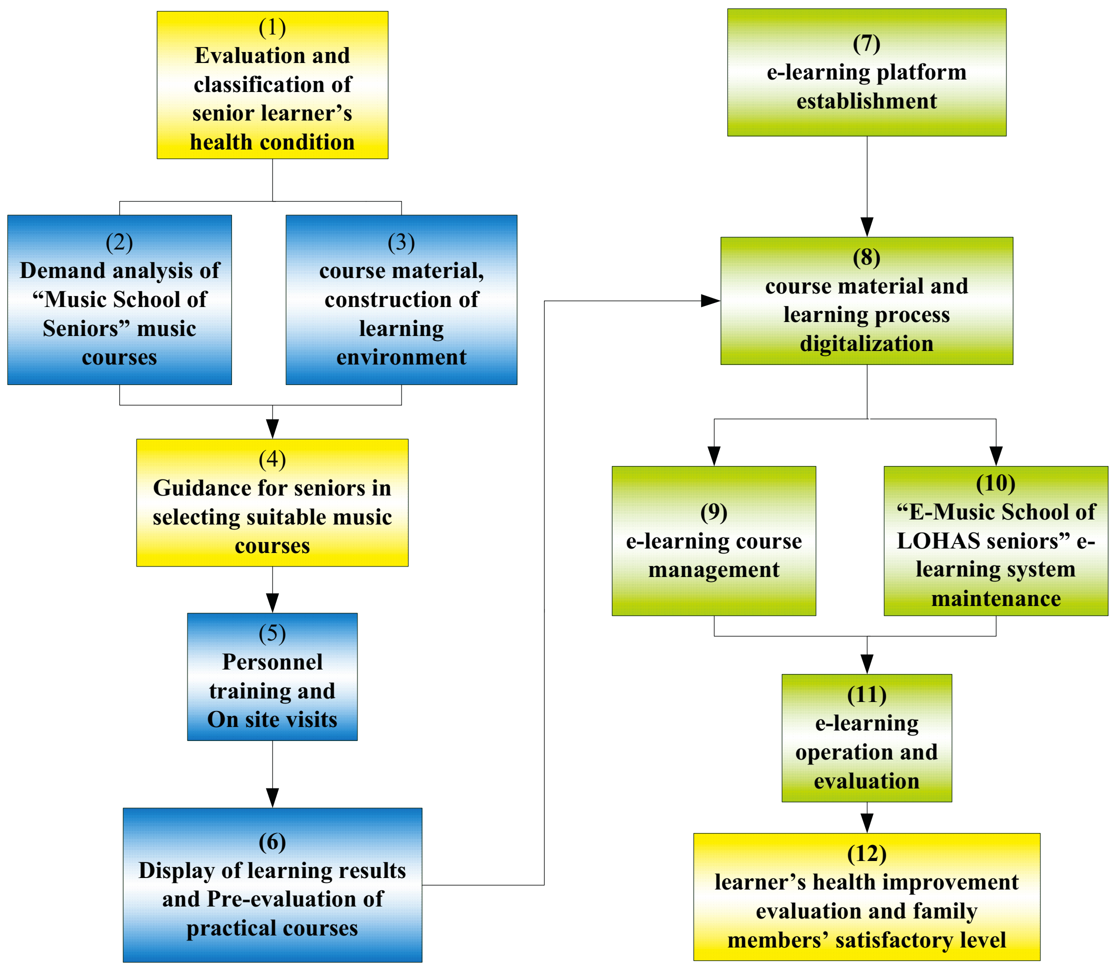


Fig 2. Twelve Steps of E-Music School of LOHAS Seniors

2.2 Steps

The project channels collaboration between Tainan National University of the Arts (TNUA) and Taiwan Association of Gerontology and Geriatrics (TAGG) and Health e-learning network (HeN) team of Ministry of Education. With a multidisciplinary collaboration, TNNUA will be responsible for the course design and teaching practice, the e-Music School website will be established by the HeN team of Ministry of Education, and TAGG will perform the evaluation of health function before and after the music learning courses. The project's goal is to improve the body and the mind of the seniors through musical education, and will be achieved by the 12 following steps:

(1) Evaluation and classification of senior learner's health condition

This step will be taken to assess the health conditions of the seniors before they participate in the courses; after putting the seniors under different health classification, the seniors will have suitable music courses issued to them. All assessing models are provided by TAGG, which will also be analyzing the results.

(2) Demand analysis of "Music School of Seniors" music courses

A feasibility research will be conducted to examine the possibility of developing industrialization of arts and humanities education for seniors in Taiwan. The data will be collected via survey. Also, a course demand survey will gather information about tastes in music and eagerness in activity participation in order to build foundation for developing course materials.

(3) Design of "Music School of Seniors" course material, construction of learning environment

The development process of the course will first be based on the field of arts and humanities education and the analyzed result from step A1-1 (course demand analysis), and then goes on to select and design course activities, music, instruments, environments. Lessons of instrument performing, singing, and music appreciation will also be designed. Another move is to categorize the functional music courses for seniors. To bring positive impact to the healthy aging process, the learning environment will also be ensured with security, academicism, and entertainment.

(4) Guidance for seniors in selecting suitable music courses

The project has rich course materials that allow the senior learners to choose favorable classes according to their Health Classification, taste, and hobbies. For those people who are interested in the courses, orientations will be held to explain the goal and contents of the course, and assist the future learners choose suitable courses. A single course will be divided into three sessions; there are eight weeks in a session. Students attend the class weekly, and take two classes on class day. Each class lasts for 45 minutes.

(5) Personnel training and on site visits

The project will recruit at least 15 instructors to design the lessons, design the teaching plans, conduct classes, report of course materials, and participate in researches. Three technicians will be hired to construct and maintain the e-learning system. Lastly, a number of six or more administration assistants will take the job of communications and administrative management.

Starting with the training of instructors, they will be informed with their rights and responsibilities, emphasis on the designs of course material and the teaching process. Furthermore, the project planned to hire experts to train the new coming instructors on skills for teaching senior music courses. After the training, new coming instructors will proceed in teaching, and complete teaching journals and reports. During the teaching process, educational experts and scholars will be taking visits to the teaching site and counsel the instructors with better teaching strategies; in the meantime, they will develop strategies on assessing the teaching process.

With the development of this project, we provide our target learners opportunities to learn; with our experience in executing the project, we provide department of music in other universities a guide for developing their own senior music courses, giving job and skill sharpening opportunities to music major students. Music education industry can also be benefited by acquiring the methods to develop senior music courses and raise their competitiveness. The project has in mind collaboration with the medical care industry. By uploading music courses on the web, home bound seniors could have a chance to learn music, in order to uplift health condition and life quality. The plan is to work with related arts and humanities industry like Angels Art Gallery in Taipei, Taiwan and create products that combine the characteristics of music and art.

(6) Display of learning results and Satisfactory level of practical courses

After each part of the course, the project will hold an exhibition on displaying the results of "Music School of Seniors", and a survey on the course satisfactory level; hired assistants will be assisting the learners on filling out the surveys. The survey of satisfactory level will impart us if our courses are in accord with the needs of the seniors, and stands as one of the grading for instructor's evaluation.

(7) E-learning platform establishment

The project obtain a vision of digitalized course materials, providing the learners an e-learning system for online learning; e-learning works as a way to share learning experiences with others, raise motivations, and results of learning. The online courses will be constructed based on the practical courses.

(8) Course material digitalization and learning process digitalization

The digitalization of practical materials of "e-Music School of LOHAS seniors" will provide the students a different learning method, and the teachers a medium to draw reference or store materials for the courses. Besides practical materials, the Learning Results Display Ceremony and other activities will be posted on the web, and allows the instructors and the students in online sharing; the family members and friends will easily share the learn accomplished achievement of each learner. Digitalization served as a function in preserving the course materials and promoting the project.

(9) E-learning course management

After the construction of online e-learning system, managing the courses will be the next step. One goal of e-learning system is to assist the instructors and students to get involved in e-learning. The system allows the instructors to upload teaching plans, and work as an e-portfolio to preserve and display student's works. Besides the function of preserving and broadcasting the learning results, e-learning system can be used as reference for future researches.

(10) "E-Music School of LOHAS seniors" e-learning system maintenance

The online e-learning system will continue to be maintained and used, and will be accessed by the instructors and the students; it will also be used to promote the project itself.

(11) E-learning operation and evaluation

A course material management and uploading crew will manage the online punch clock in order to take attendance of the instructors. The volunteers at the senior care centers will assist the students in browsing for courses, remind the students to attend classes on time, help seniors capable to look up for course information, and inform the family members to watch the student's learning progress and give encouragement to the students. The project will implement a satisfactory level survey for the e-learning system in mid-session (A month after each course session started) and in the end of the session, in order to get feedbacks on whether the system is user-friendly, or if the course contents are clear and frequently used; these feedbacks are the basis for future amendments of the system.

(12) Learner's health improvement evaluation and family members' satisfactory level

The central idea of the project is to improve the physical and mental health of the seniors through musical education. After every session, an investigation of senior daily life satisfactory level will be held. The investigation includes measurements of physical indices, mental health survey, and interviews of focal groups. These measurements can discover the changes in learners' physical and mental health condition, and examine the satisfactory level of the student's family members; after the investigation, the observation can determine the effectiveness of the course, and suggest future developments of related industries.

3. Results

The multidisciplinary research team has reached an agreement with Taipei city Zhao-Zu Residential and Nursing Home and Wenshan Elderly Service Center to provide regular music courses on site. After digitalizing the course materials and video-recording all teaching, demonstration and practicing by students with a Powercam authoring tool. The digital contents will be edited and uploaded to e-Music School (a XMS e-learning platform) by HeN team to provide an interactive environment of network community for collaborative learning among students, and facilitation or feedback between teachers and students. It also provides an environment for family member of residents of nursing home to continue follow up the life their loved ones. The family members can instantly send their regards and appreciations to the mailbox of elderly people or to the volunteers who can help to forward their messages.

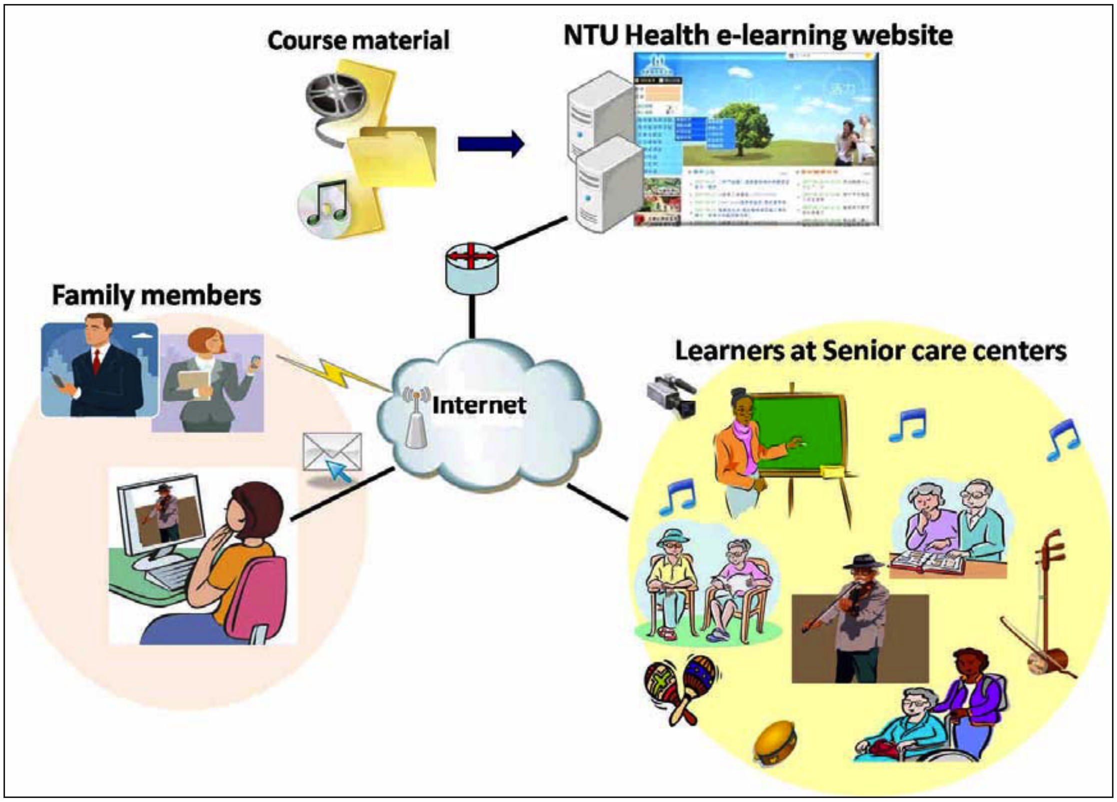


Fig3. Scenario of E-Music School of LOHAS Seniors

4. Discussion

Through combining information technology and senior medical care, the project expands musical education to the emerging business of senior care. The online e-learning system for seniors innovate a new business construction and promotion in the service industry.

Through the execution of the project, the powers of learning and researching are combined to create systematic service products and strategic service industries. Through the project, the personnel needed for related proposal planning, course development, course teaching, and feasibility analysis will be trained; providing human resources for expanding Taiwan's strategic service industries.

Carrying out the investigation and analysis of feasibility and demand, the senior satisfactory level pre and post-execution of the project, the effectiveness of the e-learning system will be presented on "E-Learning Music School of LOHAS Seniors" e-learning system demonstration website

The project creates new industries aiming at improving senior health through musical education. This brought along an immense amount of job opportunities and an effective drive for economic growth in Taiwan. Through the assessment of body functions, daily living functions, mental conditions, the project developed a set of senior courses that combines musical education and senior medical care; the courses aim specifically at improving different health functions, and creating new service products.

With e-Music School of LOHAS Seniors, it is expected to establish a e-learning platform can be shared by elderly people with their family members no matter the barriers of time of space. Family members can instantly send their feedback, appreciation to show their love to the elderly people to keep their bonds of affections not to diminish. In this project the multidisciplinary research team will combine a real music course with an e-learning environment to realize a

Web 2.0 concept. This would be the first time to be done on the music learning among elderly people in Taiwan and expected to start a new page in a comprehensive care and wellbeing for the elderly.

5. References

〔1〕National Health Research Institute, Institute of Population Health Sciences. Division of Gerontology Research, (n.d.). *Introduction.* Retrieved April 30, 2009, http://www.nhri.org.tw

〔2〕Lee, L. H., Lai, H. L., Shaw, C. K., & Chung, H. H. (2005). Effects of music on depression of the Community -residing Elders. *Tzu Chi Nursing Journal, 4*(2), 27-36.

〔3〕Hays, T. & Minichiello, V. (2005). The meaning of music in the lives of older people: A qualitative study. *Psychology of Music 33(4),* 437-451.

〔4〕Lai. H. L., & Good, M. (2002). An overview of music therapy. *The Journal of Nursing，49(2)*，80-84.

〔5〕Chen HL, Yang LL, Luh JJ, Lin GY, Establishing a National Health e-Learning Portal in Taiwan. 2006 International Conference on Distributed Multimedia Systems. Grand Canyon, 2006

Sakai 3: A New Direction for an Open Source Academic Learning and Collaboration Platform

Michael Korcuska[1]
[1]Sakai Foundation, Ann Arbor, Michigan.

Abstract— **The Sakai Project is an open source teaching, learning and academic collaboration platform built by and for educational institutions. The Sakai Project began in 2004 when Stanford, Michigan, Indiana, MIT and Berkeley began building a common Courseware Management System rather than continuing their homegrown systems or licensing software from a commercial vendor. The Mellon Foundation provided initial funding for the project. These universities recognized that research collaboration would be as important as teaching applications and developed a Collaboration and Learning Environment (CLE) that scales across many kinds of academic uses. Today, the Sakai CLE is the enterprise system of choice at nearly 200 institutions, in production settings ranging from 200 to 200,000 users. While Sakai has been extremely successful, the community has recognized a need to create a new system from the ground up. This article presents the ideas and progress on a new version of software, Sakai 3.**

Keywords—**Sakai, E-Learning, Distance Learning.**

I. Introduction

The Sakai Collaboration and Learning Environment, from versions 2.4 through the current 2.6, has been successful in enabling campuses around the world to benefit from the community source collaboration envisioned by the founding universities and the Mellon Foundation. Still, many things have changed since Sakai's inception. Sakai end users, increasing familiar with "Web 2.0" technology, are demanding an environment that is more flexible and affords them greater control. Social networking has exploded on the scene. Uses of Sakai in research and administrative collaboration have proven extremely valuable. New standards and open source projects have emerged that Sakai can leverage and integrate with. New models of web development have emerged, models that leverage client-side technology, significantly improve productivity and lower the bar for meaningful contributions and Service Oriented Architecture has emerged as a design and deployment preference for institutional systems. Most significantly, the traditional role of Course Management Systems & e-Portfolios is rapidly changing and there is broad recognition that the current platforms need to evolve substantially to meet the long-term needs of users and institutions.

The Sakai community has also learned a great deal. Increased adoption has revealed the breadth and complexity of use cases and it has become increasingly clear that portions of the code would benefit from a substantial rewrite that lowers maintenance costs while retaining flexibility to meet local needs. The limitations of the "site" as the organizing principle of Sakai are increasingly felt as institutions use Sakai in more contexts and across many years. Areas of production stress in the code and database have been identified and substantially improved, but we are reaching the point of diminishing returns with the current architecture. Sakai's SOA implementation has proven extremely valuable in practice, and yet could be improved to adhere to current standards and make it maximally compatible with new projects like Kuali. And, finally, the relative scarcity of Java developers on campus makes it imperative to simplify Sakai programming, thereby opening Sakai to a broader group of developers and increasing our capacity to innovate as a community.

In summary, the ambition is not merely an incremental improvement of Sakai 2 nor is it to copy Google. Our goal is not simply to create a better and cheaper version of Blackboard. It is time to arrive at a clearer understanding of the capabilities that represent needs unique to education and for the Sakai community to focus its development effort on providing these capabilities while taking advantage of established open-source efforts to provide more generic capabilities. We should, in short, strive to create a different type of academic collaboration system. Institutions that choose Sakai 3 will be choosing to run a qualitatively different type of system. This is the kind choice we should provide to the educational community. Not just a choice between open source and proprietary.

II. Sakai 3 for End Users

There are a number of changes envisioned in the Sakai 3.0 user experience. From a look and feel perspective the current UX Initiative is already pointing the way to a more responsive, flexible, widget-based user experience (see some samples of this work at http://3akai.sakaiproject.org). This work is resulting in significant improvements to the user experience, improvements that are critically important, but perhaps not sufficient. Beyond this usability work there is a set of conceptual changes that need to be made. Many of these changes require changes to the core Sakai architecture. We highlight a few of these changes here: Social Networking, Content Creation & Management, Moving beyond Sites and Breaking Tool Silos.

A. Social Networking

Academic research and teaching are sometimes solitary experiences, but increasingly they are becoming collaborative endeavors. There is a trend towards greater openness in university teaching, and group activity often enhances learning. The emergence of 'Social Networking' web sites such as FaceBook, LinkedIn and MySpace has created a new standard of convenience for creating online spaces that can be used to collaborate in small groups and to present profile information to peers. Innovative features such as 'activity feeds' are proving addictive in sustaining online engagement and there is increasing openness of the social networking platforms through the Facebook APIs or the OpenSocial APIs promoted by Google and is being adopted by almost all social networking sites.

However, many social networking sites require the member to grant the site owner liberal licenses covering the member's work, thus limiting the security with which confidential research or teaching can be carried out on such platforms. Moreover, the interfaces and affordances of such sites are not well adapted to academic purposes (e.g. LinkedIn profiles do not readily display publication lists).

The incorporation of Social Networking into Sakai, using the Apache Shindig project, will enable new models of interaction among users of the Collaboration and Learning Environment, in a manner suitable for academic work, but will also facilitate collaboration among Sakai institutions in which the members of a network at a trusted partner institution can be given access to the network(s) of a Sakai adopting school for research, learning, and the formation of peer groups of many kinds. While we do not know exactly what direction this work will take we believe it is critical that universities play a leading role in the development of social networking technology on campus. We cannot afford to leave this to Facebook or MySpace or Blackboard.

B. Content Creation and Organization

Authoring content inside Sakai can be awkward. There are some special purpose tools, like Melete/Modules, for creating structured content for a particular purpose. There is no tool designed to let users create unstructured or semi-structured web pages to share with others. Creating an HTML page in resources and then adding the web content tool and pointing to the HTML page is one way to go about it, but it is awkward and overly complex. Even creating references to existing content is complex and unwieldy.

And yet, creating content is a lot of what academic work is about. Instructors create syllabi for their students. Students, working alone or in groups, complete homework assignments for submission to the instructor. Research groups share and elaborate ideas in wikis and other collaborative writing software. Administrators write policy and procedure documents. And so on. In the meantime, web-based collaborative authoring tools like Google Docs/Sites/Groups have increased expectations about what is possible online.

Sakai 3 recognizes that content creation and organization is a primary activity of Sakai users, whether they be instructors or students or researchers or staff. Providing simple template-based authoring and flexible tools for organizing and presenting content will be a primary focus of Sakai 3. On the technology side, we will leverage industry standards (JCR) and open source technologies (Apache Jackrabbit) to support content storage. This will provide a significant improvement in capability (e.g. versioning) and reduce the amount of code the Sakai community needs to support. And, of course, we are not talking about creating our own HTML editor. There are a few capable open source tools we can lean on for this purpose, the current leading candidate being TinyMCE because of it's flexibility, extensibility and focus on accessibility.

Finally, Sakai 3 recognizes that many things should be treated as content. Discussion forum and blog posts, assignment submissions, user profile information and online test answers could all be usefully regarded as pieces of content. Taking an "everything is content" approach to Sakai 3 will allow much more flexibility in searching, organizing, tagging and otherwise manipulating items in Sakai.

C. Moving Beyond Sites

Sites are the primary organizing principle of today's Sakai. Site context is a deep and rigid assumption for nearly all functionality, and it stands in the way of activities that might extend across sites or operate independently of them. The notion of a group in Sakai stands as a particularly strong example of unnecessary site dependence. In Sakai 2, groups only exist within a site. If you want to address a particular group of individuals they need to all be members of a single site. If the same group needs access to multiple sites then that group needs to be recreated. In Sakai 3 groups are treated as first class citizens. Users will be capable of managing groups independent of sites. They can create groups, referencing an external system as needed (e.g. an SIS like Banner perhaps through the IMS Enterprise Specification or via a more generic LDAP provider), and later worry about what that group has access to.

This line of thinking will be applied to other items in Sakai, including Users and Content (and Tags and Permissions). Content is another excellent example where the primacy of the site is more tyranny than convenience. While content can be made public, it exists inside a particular site context. Moving content between sites or referencing content from another site is cumbersome and unnatural. In Sakai 3 content will be a first class citizen as well. Content owners can organize content in a variety of ways and make it available to various users, groups and sites as they see fit. Instead of the site's content we will think about the user's content or a group's content, both content they own and content they have access to.

A. Social Networking

Tools are another important target for change. Sakai's "tool silos" are well recognized and restrict natural workflows that reflect user models of academic activity. While more intuitive "cross tool" interactions are more and more common in Sakai 2, the underlying architecture and original technologies makes this difficult at best. Sakai 3 will be constructed around smaller units of capability (in the form of true SOA services) that can be quickly stitched together to provide intuitive workflows. We need to think of the relevant items in Sakai (from discussion posts to assignment submissions to test questions to portfolio reflections) and the activities relevant to those items. These items and activities will surface in many different places depending on the context. Sakai 3 should respect the context and present the workflows that make sense inside them.

III. Sakai for Technologists

We have learned a great deal in the last several years and, at the same time, new technologies and techniques have emerged. Much progress has been made, to be sure, on the current architecture and code base but we are reaching the point of diminishing returns. Armed with the insight gained from the current experience, we are in a good position to establish backing technologies that both improve the product and reduce the maintenance burden on the community. The resulting Sakai 3 will deliver a variety of technical benefits, including the following:

1) Scalability and Resilience: While Sakai has achieved impressive levels of scalability already with installations of over 200,000 users, achieving good performance at this scale has required a significant investment. With the knowledge we've gained about Sakai usage patterns and the inclusion of new "internet scale" open-source technologies, Sakai 3 will both achieve new levels of scalability and will make it simpler to run a smaller installation.

2) Improving Developer Productivity: Recent community efforts in client-side development have underscored just much more efficient Sakai development can be. Today, a substantial depth of technical skill and attention to a large number of details are required to achieve even results, and the consequences of this difficulty ripple outward: less gets done, fewer people are able to do it, and usability experts are effectively held at arm's length. Sakai 3 services will provide JSON data feeds allowing JavaScript developers to create user interfaces and, if desired, generally function independently of Java developers. This also frees Java developers to spend less time on user interface rendering and focus on the scalability and quality of the core services

3) Code Quality and Maintainability: The maturity of other open source projects now allows us to consider swapping out whole regions of Sakai services with 3rd-party code. By judicious incorporation of such services our overall quality can be improved, our APIs can be made more standard, and our maintenance burdens and risks lightened. In addition, Sakai has traditionally had very little test code. This oversight increases burdens on QA cycles by not exposing problems quickly enough, ultimately forcing us into a post-release reactive posture with regard to critical bugs. 3rd party code must be examined for the quality of the test coverage it brings with it and we should extend full coverage to Sakai services.

4) Installation and Maintenance: Sakai is too difficult to build for staff that are not experienced Java developers and too difficult to install for staff that are not sufficiently technical. A smaller, tighter kernel and a mechanism for easily adding/removing tools will allow new Sakai users, deployers and developers have a more positive and productive initial experience. This is important to the growth and overall health of the community. A more efficient, easier to maintain installation has many benefits, most notably that staff can spend more time on innovation.

III. Sakai Use Cases

Up to this point in time, Sakai has been defined by the existing product category that best matches how particular campuses are using it. Thus Sakai may be an ePortfolio system, a Course Management System, a Virtual Research Environment, or even a web-based file store. This has been both a strength, demonstrating the platform's flexibility, and a weakness, creating more diffuse communication and development effort. The development of Sakai 3 will allow clear delineation of the underlying platform technology and the various deployment profiles supported on that platform.

When we do what has been described above, Sakai will acquire utility in many corners of the institution beyond the current teaching and learning base. One goal of Sakai 3 is to support more open teaching and learning practices. In doing so, Sakai begins to acquire some of the characteristics of a content management system. This is explicitly recognized in some of the underlying technology choices of Sakai 3. The boundaries for Sakai will continue to blur and to increasingly be expressed in terms of which broad capabilities do we want to offer to which users rather than which product category does Sakai occupy. So rather than saying that "Sakai is a CMS" or "Sakai has an ePortfolio tool" we should say: We're using Sakai to support student portfolios and to provide technology support for teaching & learning on campus.

On the other hand, for those institutions that do want to offer Sakai as a niche it should be easy to configure an installation with those characteristics. Two examples follow:

1) Sakai 3 as a portfolio system. Content in Sakai 3 will no longer be tied to sites. The ability to make content public or private and in general, to control the access various groups have to particular content, will be much more flexible and controllable by the user. The unstructured and semi-structured content authoring in Sakai 3 will allow for free-form portfolios. Tagging will be built in, with tag vocabularies built by the user or provided by the institution, allowing content to be more easily assembled. Through all of these improvements it should be possible to create portfolios that more easily meet the typical portfolio uses cases: student self-expression, an electronic resume and institutional analysis.

2) Sakai as a group collaboration system. Group projects are often more fluid than classes and definitely last longer. Participants can change rapidly, participants from outside the campus are more likely to be involved and, generally speaking, the proportion of users creating content for use by the group is higher than in courses. Research projects may have special archiving needs. Projects also often want to have a public presence and be able to use the "same space" for the work internal to the project and communication with the outside world. Sakai 3's group and content management capabilities should provide an attractive package for organizations wanting an on-campus system for project collaboration. Yet Sakai 3 should integrate will with off-campus systems (e.g Google docs) and therefore provide individual groups with flexibility on the specific tools they use.

III. Summary

In summary, the Sakai 3 effort is designed to bring a number of benefits to campuses, including:

1) Increased user satisfaction – A more fluid and flexible Sakai, one that is both pleasurable and efficient to work with, will allow users to focus energy on improving the quality of their work.

2) Improved stability & quality – In addition to the obvious user satisfaction that comes with stability and quality, your campus IT and support staff will have more time for other activities.

3) Increased scalability – The ability to support more users per application server will reduce the overall cost of ownership of Sakai and can let your organization serve additional customers (e.g. a local school district) that might previously have been too expensive to consider.

4) Fewer local customizations – As an Open Source product, customizing Sakai for your local needs should continue to be one of the main benefits of using Sakai. By ensuring that more use cases are covered "out of the box," however, your local customizations can really focus on what is unique to your organization.

5) Simpler integration – Sakai already has a reputation for being an excellent application to integrate with other campus systems. By building on and improving Sakai's service-orientated architecture, these benefits will continue to accrue as more campus systems support SOA.

6) Easier development – Allowing a wider variety of developers to contribute to Sakai creates a wonderful virtuous circle. More participants can help the project and existing contributors will be more efficient. This means more staff time for local customizations and, more importantly, innovation.

All of these add up to a lower total cost of ownership and better value for your investment. What your campus chooses to do with the additional resources—spend them on other project or increase the amount of innovation in the development and use of Sakai is up to you.

Acknowledgments

I would like to thank the Sakai Community for their contributions to this document, most notably John Norman, Ian Boston and Clay Fenlason.

Address of the corresponding author:

Michael Korcuska
Executive Director
Sakai Foundation
3057 Hillegass Avenu
Berkeley, CA 94705
Phone: +1-510-599-2596

Concept Map Supported E-Learning Implemented on Knowledge Portal Systems

Jyh-Da Wei
Department of Computer Science and Information Engineering,
Chang-Gung University, Tao-Yuan, Taiwan

Tai-Yu Chen*, Tsai-Yeh Tung* and D. T. Lee*
Department of Computer Science and Information Engineering,
National Taiwan University, Taipei, Taiwan

Abstract

E-learning techniques have grown quickly nowadays to complement the conventional learning system. In this research field, integration of knowledge management (KM) and e-learning has been a tendency that coordinates the creation, organization, and delivery of a learner-oriented knowledge management system. Knowledge portal systems are usually efficient vehicles to realize this integration for we can create new content types on the portal sites to serve e-learning purpose. Nevertheless, a knowledge portal site allows the members to create and share their knowledge objects; implying that the website containing group knowledge grows itself unlimitedly and unexpectedly. To display the sitemap of a knowledge portal and observe the concept map of group knowledge, we have developed a visualization tool referred to as the Site-And-Concept (SAC) map generator. In this paper, we apply this map generator on a knowledge-portal-based e-learning system. Users can convert the generated concept maps into SCORM-compliant course packages with sequencing and navigation supports. They can further edit the learning content and share it to peer learners, and thus establish a collaborative learning environment.

Keywords:
Knowledge management, e-learning, knowledge portal, concept map, SCORM, collaborative learning

1 Introduction

Integration of knowledge management (KM) and e-learning has been an attractive research topic for the many dependencies and interrelations between these two fields [10, 11]. Learners over the computer networks usually benefit from KM, whereas the research within KM also considers learning as an import part of knowledge sharing processes. The integration therefore is of high potential to complete knowledge creation and delivery cycle. In the beginning of the research into integration, the KM systems are used as online materials supplementary to regular face-to-face teaching [4, 8]. The group activity provided by KM systems was immediately found highly effective for the traditional classroom teaching [1], and thus the research towards integration of KM and e-learning receives increasing attention ever since.

Nowadays, most popular integrated models tend to coordinate techniques following different perspectives of KM and e-learning [11, 12]. KM is mainly related to an "organization" perspective and can be seen as the process of integrating information and making it explicit, organized, accessible, and portable [10]. On the other hand, e-learning derives a "regeneration" perspective. The perspective of e-learning can be further considered involving short-term performance and long-term development [5]. For the short-term performance, the user learns the target materials just in time and in context; whereas in the case of long-term development, the learner can feedback the knowledge base by producing new knowledge objects or supporting behavior and tendency analysis.

Knowledge portal systems are usually considered as efficient vehicles to incorporate knowledge management with e-learning functions. A knowledge portal is a kind of website aimed to create, organize, and store explicit knowledge for easy access, sharing and manipulation. Furthermore, research of education [3, 14] has developed several manners to convert knowledge objects into learning assets and to facilitate packaging of content aggregation. These methods are convenient to implement and have proved practical; however, there still exist performance issues that deserve further research. For example, we may suffer the *cogni-*

*Authors are also with the Institute of Information Science, Academia Sinica, Taipei, Taiwan.

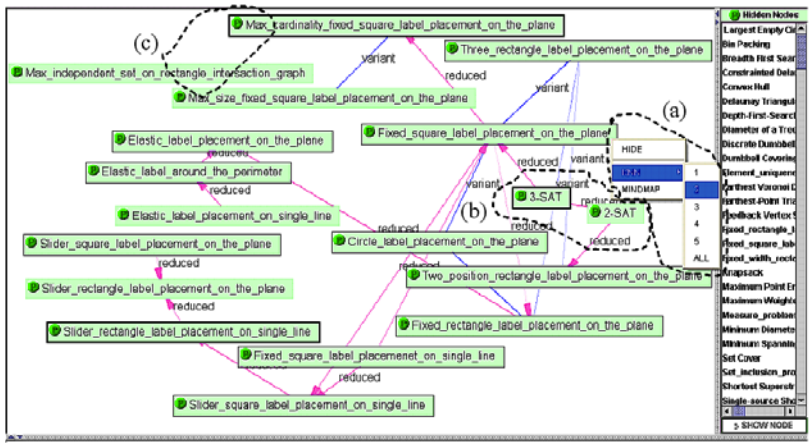


Figure 1. Showcases of the Site-And-Concept map. (a) Right-click k-NN function; (b) Aggressive learners can display k nearest neighbors of a target object; (c) Expert users may be interested in verifying the completeness of group knowledge.

tive disconnection problem [5] if we create e-learning content types on a portal system immediately. The e-learning function may has limited performance because the isolated learning content cannot cooperate with the underlying knowledge ontology smoothly. In this paper, we present our previous developed Site-And-Concept (SAC) map generator, which constructs concept map linking knowledge objects and thus helps users visualize the formation of group knowledge. Then, we use this map generator as a mediate tool to produce SCORM-compliant course packages with supports of sequencing and navigation. As for the practice platform of this work, we choose a locally developed knowledge portal, called Open Computational Problem Solving (OpenCPS, http://www.opencps.org) Knowledge Portal.

2 Map Generator to Visualize the Formation of Group Knowledge

In comparison with conventional websites, knowledge portals have three features: (1) The contents are contributed by all the portal members, not just by the web design community. Accordingly, the contents of a knowledge portal can be regarded as group knowledge, which aggregates all the intelligent assets from the userbase; (2) Apart from an axiom ontology given by the portal designer, no detailed sitemaps or blueprints are prepared to outline the contents in advance, i.e., the contents of the website grow in an unplanned and unpredictable manner; (3) The developed knowledge objects are associated with conceptual elements defined within the ontology of the knowledge portal. Consequently, we can construct the sitemap of the portal site by adopting the notion of a concept map. In consideration of the above three aspects, we developed a visualization tool in our previous work [6, 13] for constructing a Site-And-Concept map to improve the quality of a knowledge portal.

Figure 1 shows the experimental results conducted on our OpenCPS website. The typical operations of the SAC map generator can be categorized into three classes, i.e., the basic, learner-mode, and expert-mode operations. To simplify the demonstration, the knowledge objects only include the problem objects inside the OpenCPS portal. The relations are labeled as “reduced-to”, “equivalent”, “variant” and so on.

1. **Basic operation:** The basic search operation is to use keywords to search the target objects. For example, users may type “map labeling” as the keyword and thus draw a SAC map as shown in Fig. 1 except parts (b) and (c). They can also right-click the relation links to look up and edit the property of the relationship of interests.

2. **Learner-mode operation:** For an aggressive learner, to show the knowledge objects associated with given keywords is not enough. For example, the 3-SAT and 2-SAT problems cannot be found when we search for keyword “map labeling”, but these two problems are important for a student to get a thorough understanding of how difficult map labeling problems are, because they can be polynomially reduced to the “Fixed square label placement on the plane” problem. The right-click k-NN function helps display k nearest neighbors of a target object, as Fig 1(b) shows.

3. **Expert-mode operation:** Expert users are usually in-

Figure 2. Complete map of current knowledge objects on the OpenCPS website . We do not apply any filtering and layout process so that this picture may simulate the image of human cognition without concept map supports.

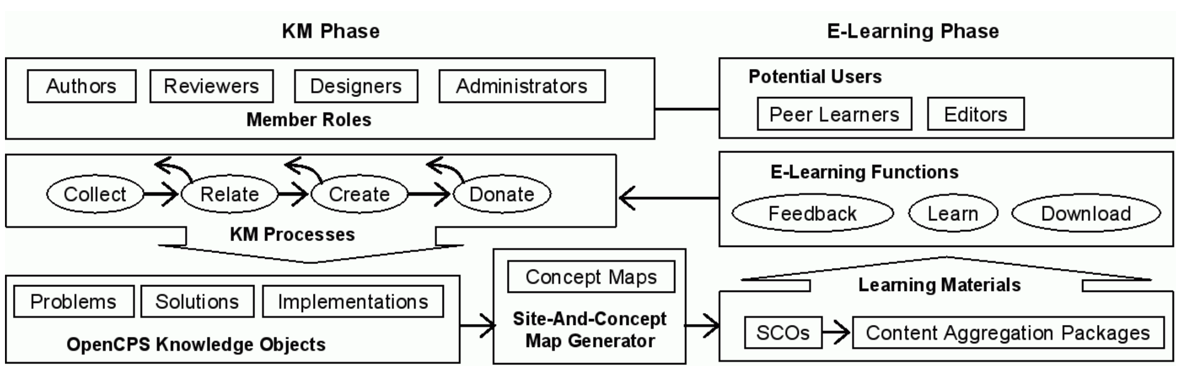
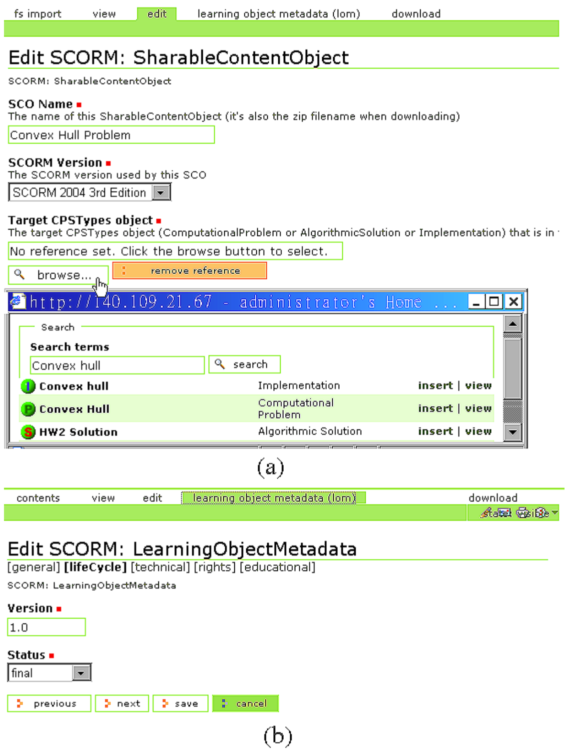


Figure 3. Architecture to integrate KM platform and e-learning processes by constructing group knowledge concept maps. This integration method using SCORM-compliant content types implies the implementation of an emerging educational perspective, "learning by teaching".

terested in verifying the completeness of other members' knowledge. The SAC map generator is suitable for the experts to make contributions. If they are very certain that one problem object is related to another and this relation cannot be found by the k-NN function, they can use the keyword search again. The search result will imply that they either create such a new problem or add a relation link between these problems. For example, an expert user firmly believes that the "Max cardinality fixed square label placement on the plane" problem is a variant of the "Max independent set on rectangle intersection graph" problem, and a follow-up search also reveals that the latter problem object is an isolated node as shown in Fig 1(c). Accordingly, this expert user could generate such a relation in OpenCPS and label the relation link, and thus make the knowledge structure more complete.

The SAC map generator allows the members to examine the current knowledge formation by visualization. Thereby, the group knowledge of this knowledge portal can thus serve as an aggregation of content for learning purposes and get enhanced from users' feedback. As a comparison, Figure 2 is the complete map of the knowledge concepts captured when we do not apply any filtering or layout processes. This picture simulates the image of human cognition without concept map supports. This shows a rather messy situation. We can hardly make out any of the concept relations, much less to examine or learn from this knowledge portal. Therefore, we believe the map generator can prop-



(a) (b)

Figure 4. SCORM-Compliant Content Objects converted from a given concept map. (a) a “SharableContentObject" (SCO) on opening the “edit" tab and the list of SCOs collected from the corresponding concept map; (b) editing the SCO’s metadata.

erly enhance the performance of a knowledge portal.

3 Knowledge-Portal-Based E-learning System with Concept Map Support

Figure 3 shows the architecture for integrating the concept map generator into a knowledge-portal-based e-learning system. In this work, we follow the SCORM (Sharable Content Object Reference Model) standard to convert a generated concept map to a course package including the sharable content objects (SCOs) and the properties for sequencing and navigation. To suggest users an appropriate learning sequence, we take account of near optimal solution algorithms to the ATSP (Asymmetric Traveling salesman problems) such as the shortest-edge-first algorithm accompanying with the 2-opt advancement.

In our design, the concept maps are generated automatically by assessing the knowledge relationships based on the probability of “shared referenced papers”. Assuming that two knowledge objects (i.e., portal pages and nodes of the maps) M and N have $Q(M)$ and $Q(N)$ referenced papers in the citation database respectively and the number of shared referenced papers is $Q(M,N) = Q(N,M)$, we define the cross-reference probabilities $p(M,N) = Q(M,N)/Q(N)$ and $p(N,M) = Q(N,M)/Q(M)$. The map generator thereby indicates a link between the two objects if $max(p(N,M), p(M,N)) \geq T_p$, where T_p is a threshold for mining this association rule. Notably, the default arrow direction on the concept map is from N to M if $p(M,N) > p(N,M)$ and vice versa. The probability is now also concerned to count the edge weight, i.e., $cost(M,N) = 1/p(M,N)$, for use in this work. Accordingly, we approximate the solution for the ATSP corresponding to the given concept map, and then figure out some learning sequences of the learning package.

On converting a concept map to a course package, OpenCPS switches to an edit form, as shown in Fig 4(a), for the user to input the SCO names, select the SCORM version, and modify the suggested learning sequence if necessary. For each SCO, the user can also click “learning object metadata (lom)” tab next to the “edit” tab on the same page to input the metadata. Figure 4(b) shows that the subpage within the “lom” tab has deeper tabs binding to “general”, “lifecycle”, “technical”, “rights”, and “educational” categories, for which the user can fill out all the associated metadata. This work adopts the SCORM standard, because SCORM clearly defines the format of content object metadata and content aggregation manifest files, and can therefore facilitate us in creating new content types with relevant data columns. Content types following the SCORM standard also inherit its benefits, i.e., they can be accessible, adaptable, affordable, durable, interoperable, and reusable.

4 Conclusion

For integration of knowledge management and e-learning, Ley et al. [5] addressed that a typical workplace must link a work space, a learning space, and a knowledge space. In this case, the first problem that developers may encounter is the “cognitive disconnection”, namely, “each of the spaces has an inherent structure which mirrors to some extent mental model of the people who are using it” and it might be hard to connect available KM model to learning activities. In this paper, we develop a mechanism that generates a local concept map and then converts the knowledge nodes and relation links of interests into sharable content objects (SCOs) with supports of sequencing and navigation properties.

Users can generate their concept maps of group knowledge and then manipulate the objects thereof instead of working directly on the substantial objects which belong to the knowledge portal. This implementation evades the problem of cognitive disconnection by using the mediate concept maps, and can seamlessly support learning activities with the ability to organize information and knowledge resources on the website. Additionally, users are allowed to transmit their packages to teaching assistants and all peer learners. This mechanism can therefore also realize the aspect of learning by teaching [2, 7, 9].

Acknowledgment

This work was supported in part by the National Science Council under the Grant NSC98-2218-E-182-005 and the Office of Research and Development of Chang-Gung University under Grant UERPD280041.

References

[1] R. Benbunan-Fich and J. B. Arbaugh. Separating the effects of knowledge construction and group collaboration in learning outcomes of web-based courses. *Information & Management*, 48:778–793, 2006.

[2] E. Ching, C.-T. Chen, C.-Y. Chou, Y.-C. Deng, and T.-W. Chan. A pilot study of computer supported learning by constructing instruction notes and peer expository instruction. In *Proc. 2005 conference on Computer support for collaborative learning*, pages 63–67, 2005.

[3] R. S. Choenni, R. Walker, R. Bakker, and W. Baets. E-learning as a vehicle for knowledge management. In *Proc. 14th International Conference on Applications of Prolog*, pages 379–392. 2001.

[4] D. E. Leidner and M. Fulle. Improving student learning of conceptual information: Gss supported collaborative learning vs. individual constructive learning. *Decision Support Systems*, 20(2):149–163, 1997.

[5] T. Ley, S. N. Lindstaedt, and D. Albert. Integration of e–learning and knowledge management – barriers, solutions and future issues. *Lecture Notes in Artificial Intelligence*, 3782:189–202, 2005.

[6] Y. L. Lin, J. D. Wei, G. C. Lee, and D. T. Lee. A visualization tool for the sitemap of a knowledge portal and the concept map of group knowledge. In *Proc. 5rd International Conference on Knowledge Management*, pages 179–186. 2005.

[7] S. Y. Okita and D. L. Schwartz. When observation beats doing: learning by teaching. In *Proc. 7th international conference on Learning sciences*, pages 509–515, 2006.

[8] V. V. Phoha. An interactive dynamic model for integrating knowledge management methods and knowledge sharing technology in a traditional classroom. In *Proc. 32th SIGCSE technical symposium on Computer Science Education*, pages 144–148, 2001.

[9] M. Pollock. Basic mechanics: learning by teaching - an increase in student motivation (a small scale study with technology education students). In *Proc. 35th Annual Frontiers in Education*, pages T2F/12–17, 2005.

[10] E. Ras, M. Memmel, and S. Weibelzahl. Integration of e–learning and knowledge management – barriers, solutions and future issues. *Lecture Notes in Artificial Intelligence*, 3782:155–164, 2005.

[11] N. H. J. Razzouki. Perception of e-learning and knowledge management: a trend toward integration. In *Proc. 2004 International Conference on Information and Communication Technologies*, pages 97–98. 2004.

[12] A. Schmidt. Bridging the gap between knowledge management and e–learning with context–aware corporate learning. *Lecture Notes in Artificial Intelligence*, 3782:203–213, 2005.

[13] J.-D. Wei, T.-Y. Chen, T.-Y. Tung, and D. T. Lee. A visualization tool to construct the reference-based relation map for group knowledge of a knowledge portal. *2007 IEEE International Conference on Information Reuse and Integration*, pages 304–310, 2007.

[14] J.-D. Wei, T.-Y. Tung, T.-Y. Chen, and D. T. Lee. Integration of knowledge management and e-learning – an implementation towards learning by teaching. In *Proc. 2007 International Conference on Information and Knowledge Engineering*, pages 414–420, 2007.

An Implementation of the Tools in the Open-source Sakai Collaboration and Learning Environment

Yasushi KODAMA[1], Tadashi KOMORI[2], Yoshikuni HARADA[3], Yasushi KAMBAYASHI[3], Yuji TOKIWA[4], Kazuo YANA[1]

[1]IT Research Center, Hosei University, Tokyo, Japan

[2]Aimluck Inc., Tokyo, Japan

[3]Nihon Institute of Technology, Saitama, Japan

[4]Research Center for Computing and Multimedia Studies, Hosei University, Tokyo, Japan

Abstract— **This paper describes a novel approach to implement tools in the open-source Sakai CLE that is a scalable approach to meet requests from large universities with student populations of over 10,000. We have already designed and implemented the I18N help and class attendance recording system on our campus Sakai system. Upon redesigning our tools on Sakai CLE, we have integrated the aspect-oriented approach in order to obtain flexibility and scalability.**

Keywords— **e-learning, course management system, open-source software, Sakai CLE**

I. INTRODUCTION

Recently, many institutions of higher education have introduced Learning Management Systems (LMS) and Course Management Systems (CMS) for instruction and research. We have designed and are now developing our own scalable system based on the well-known Sakai CLE.

At large universities, registrants run from thousands to tens of thousands. Since a large university offers many different courses, we have to accommodate a variety of requests as the system functionalities. Such requests include the extensibility of LMS and CMS. Our on-site educational staff should be able to modify them. In order to implement our LMS and CMS, we chose PHP and Java as the description language. PHP is a language most widely used for describing information systems, and Java is an object-oriented language suited to building large systems. We can expect they increase reusability and flexibility in our system.

We selected the open source educational support system Sakai CLE (Sakai Collaboration and Learning Environment), which uses the object-oriented language Java. Using Sakai CLE as a starting point, our effort aims to expand the system's functionality, extending the design and development of new tools. This paper focuses on our efforts in design and implementation of the system we have completed to date.

II. EXPERIENCE TO DATE

Since the year 2000, we have started extensive research on distance learning and e-learning at the Hosei University IT (Information Technology) Research Center. The same year, we also opened the Hosei University Research Institute, California, located in Silicon Valley, where we started offering distance learning from the US. We initially recorded actual courses and then offered them to students in an on-demand format for use in preparation and review. Since 2003, however, we have been improving the system to record contents specifically for on-demand distance learning. At the same time, we adapted the Sakai CLE educational support system for the basis of our distance-learning system, which has proved useful for communication between teachers and students. Since instruction was offered primarily in English in the beginning, the English language version of Sakai CLE was sufficient. Recently, however, the number of requests to use Sakai CLE for regular courses prepared in Japanese at Hosei University has increased. In response to the requests, the IT Research Center began design and development of a set of new dedicated tools for the Sakai CLE in 2006: the internationalization (I18N) and Japanization (J10N) versions. We introduced our overall approach regarding Sakai CLE in [1]; in this paper, we mainly describe implementation of a version that has been further improved.

At present, Sakai CLE is being designed and developed primarily at the Sakai Foundation. We established the Japan branch, Ja Sakai, where we engaged in disseminating the Sakai CLE.

III. DEVELOPMENT OF SAKAI CLE TOOLS

Sakai CLE activates on what is called a servlet based on Apache Tomcat. Our portal framework is based on a product developed independently at Stanford University. The system has since been improved and incorporated new frameworks such as Spring and Hibernate, and is resulted in the Sakai CLE. The current version is Sakai 2.6.1. Versions under development have also been released separately [2].

We first implemented the Sakai CLE for distance learning in English, thus we just used the English-language version. As we began using it for regular classes in Japan, we have received a number of requests to internationalize the Sakai CLE, or at least localize it for Japanese. By 2005, the Sakai CLE has already supported internationalization (mainly thanks to a Spanish implementation), but the adaptation was not complete (garbled characters appeared in some messages and so on). We first fixed the garbled characters in the messages.

In Sakai CLE, the help window can be started up from any screen by clicking the button. However, only one language type can be displayed in the help messages. For example, when the computer's locale is changed to Japanese (ja), messages are displayed only in Japanese not in English. In [1], we introduced the first version. Here we will discuss the second version, which can also handle Chinese language messages.

A. Development of a help tool using an aspect-oriented approach

In [4], they discuss the large-scale software development of e-learning system using aspect-oriented programming (AOP) and design pattern. We also redesign our tools using AOP.

The Spring and Seasar2 frameworks in Java servlets offer AOP that uses dependency injection (DI). In internationalizing Sakai CLE, introducing the same sorts of frameworks was effective in enabling the use of AOP.

We used AOP to handle the display of help screens for internationalizing them [3]. The procedure for internationalization switching under AOP was as follows:

1. Specify the Spring setup file local_langset1.xml in the servlet setup file web.xml.
2. Specify the implemented LangSet1MethodInterceptor and LangSet2MethodInterceptor in the setup file local_langset1.xml. We tried implementing LangSet1MethodInterceptor to contain a set of two languages and LangSet2MethodInterceptor to contain a set of seven languages.
3. After compiling the setup file local_langset1.xml, we applied the changes made using AOP by restarting the Apache Tomcat web container.

In internationalizing the tool, we implemented Sakai CLE so the language displayed in the help screen at the top left is retrieved by the method getLangs(). LangSet1MethodInterceptor and LangSet2MethodInterceptor are Java class files that implement org.aopalliance.intercept.MethodInterceptor. In the setup file local_langset1.xml, we use the pattern matching function of the method invoked with the regular expression, weaving it in with the method getLangs() of the existing class. When the existing method getLangs() is called, it returns a new language set (two languages when using LangSet1MethodInterceptor) as the return value. As a result, the language set that is displayed will change depending on the class used. By clicking a language, the help menu at the bottom left is displayed in the selected language.

IV. CLASS ATTENDANCE RECORDING SYSTEM

Even in large universities, the class attendances are usually managed based on paper attendance slips. By

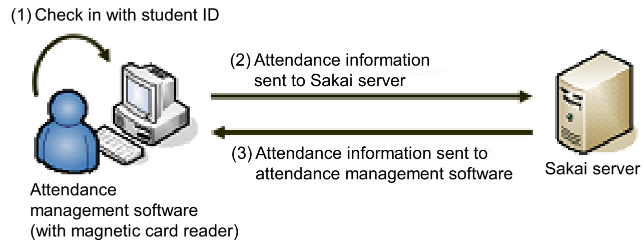


Fig 1: Class attendance recording system in overview

introducing a class attendance recording system, attendance can be managed on a network using PCs to gain efficiency and accuracy. Furthermore, registering attendance data in the Sakai CLE grade management tool (gradebook) enables total student performance management. We implemented a class attendance recording system using Sakai CLE. It consists of attendance management software (on a PC) and an attendance management tool (on the Sakai CLE server). Students can be checked in and out using their student ID by linking the attendance management software to a magnetic card reader.

Based on the assumption that the attendance management software will be used in a classroom environment where there is no network available, we decided to run it as a stand-alone tool rather than rely on it having to run on the Sakai CLE server. Recording attendance manually is time-consuming, so we provided for automatic student check-in using a magnetic card reader connected to a PC.

A. Attendance information management

The attendance management software stores the student ID number and time of student ID check-in/check-out in the database by passing a student ID through a magnetic card reader upon arrival and departure (entering and leaving the room, two times total). If the student ID number is not registered in the attendance management software when the magnetic card reader reads the ID number, the student name is recorded as GUEST and the student ID number is automatically added.

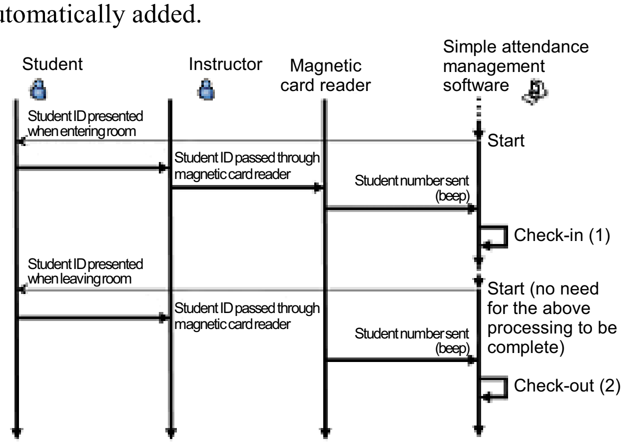


Fig. 2: Check-in/check-out flow

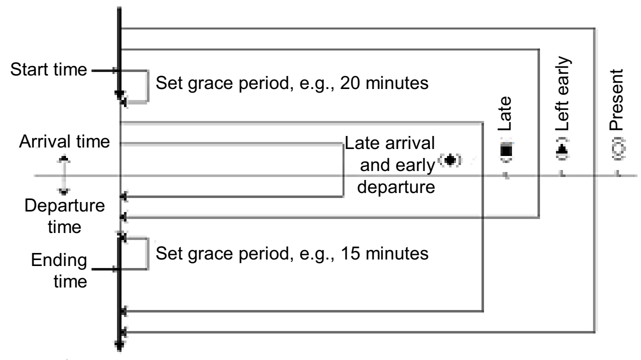


Fig. 3: Calculating attendance status for students

After the end of the lecture, or if the magnetic card reader could not read a scanned ID, additions or edits can be entered manually. In order to make it clear that the corrections were made manually, the character string (Manually input) is displayed so the operator can determine whether information was logged during the lecture. Fig. 2 shows the specific check-in flow for the attendance management software. There are a total of seven statuses pertaining to student attendance, namely Present, Absent (also includes when a student ID was not passed through the reader), Late, Left early, Late and Left early, as well as Reason for manual input and Error (when departure and arrival times not entered). The decision for the attendance status is calculated as shown in Fig. 3 based on permissible ranges for late arrival and early departure grace periods.

B. Student information management

While student information can be managed for each lecture, student information can also be edited independently. Student information managed by the attendance management software consists of student name, student ID number, and affiliation (department name). When a student's attendance has already been logged for another lecture, registered student information can be listed based on attendance information when that course name is selected. Student information can be output in Excel format for external use.

C. An implementation of the attendance management tool

We have designed and developed a new attendance management tool as a dedicated Sakai CLE tool. Upon importing attendance check-in/check-out information data collected by attendance management software into Sakai CLE, we can check the student attendance statuses on Sakai CLE (see the figure below) on a web browser. We have further extended the system so that the ratio of the attendance status (percentage) for each student can be added as a point number to the existing performance management tool (gradebook). This allows unified management of student performance - including attendance - on Sakai CLE. Below we discuss attendance information management.

D. Attendance information management

Student attendance status is displayed as a list in the attendance management window of the attendance management tool. Lecture information is read-only, with editing performed in the attendance management software (on a PC).

After clicking **Attendance Lists** on the top menu, attendance information for students who attended a lecture is displayed as a list selected from a pull-down menu. It can be displayed either in the form of symbols or points (up to 100). The window is narrow enough to be displayed on a single screen, and is organized by lecture schedule stacked vertically by date.

Clicking a student ID displays a student profile, and clicking a lecture schedule link allows editing of student arrival and departure times. After changes, the attendance list is updated again automatically to enable instant checking of recalculated attendance points.

E. Input using an IC card reader

In version 1 of the attendance management software, students checked in using a magnetic card student ID; in version 2, however, we improved this by using an IC card reader for check-ins. In this section, we describe implementation using an IC card reader.

Using IC cards is more convenient and efficient way to check in than using magnetic cards to do the same things. The Pasori (RCS-320) IC card reader is inexpensive and easy enough to implement for this purpose.

Device drivers that handle the Pasori include SDK for felica, felicalib and libpasori. We chose libpasori (free software) for this implementation. Since it performs transactions with Pasori, a USB device, with a library written in C, we rewrote it in Java to integrate in our current system.

The Pasori IC card reader can get a variety of information from the Felica IC card; in the attendance management system, we retrieved the IDm information and used it as the student ID. The procedure for fetching the IDm information is as follows.

1. Pasori is detected from a USB device.

2. Initialization is performed.
3. Polling is performed and data entry is awaited.
4. Management data is converted into a byte string.
5. Procedure is repeated from step three.

As described above, the data that is entered is placed in busy-wait status.

Fig. 4 shows the attendance management software and Pasori IC card reader.

V. DISCUSSION AND CONCLUSION

We have reported findings obtained from the design and development of dedicated tools for the open-source Sakai CLE based on the unique requirements at Hosei University. We have begun by fixing garbled characters in messages that were internationalized or converted to Japanese, and modified existing dedicated tools (help tools). We also implemented class attendance recording system. This is the unique dedicated tool for Hosei University. Furthermore, in our modifications of the help tool, we have found that we needed to provide additional descriptions for Java class files when using AOP to add new languages to help messages.

We intend to distribute the class attendance recording system as open source software like Sakai CLE. To distribute it as open source software, we must first write more generalized descriptions for those sections that are unique to Hosei University. The class attendance recording system requires a variety of expanded functionalities. In this paper, we have described how we developed new attendance management software for the PC, and how we linked the software on the PC to Sakai CLE.

In fiscal 2007, we conducted an experiment testing the class attendance recording system in actual classes. We found it had three major benefits:

1. Reduction in late arrivals and early departures (increased motivation to attend).

Fig. 4:The attendance management software and Pasori IC card reader

2. Student awareness of their attendance records.
3. Increased use of the course management system (Sakai CLE).

For point 1 in particular, students became aware of how rigorous the automated management of attendance times had become, and the benefits were quite significant. Points 2 and 3 could be considered as side effects of linking to a course management system. We will conduct further experiments to verify the advantages we have obtained; along with a detailed analysis (including the use of questionnaires) so that we can continue to improve the functionalities.

We are now looking into the design and development of a tool for use with on-demand video lectures as a new dedicated tool development project. To provide on-demand video lectures, we need to collect statistical data such as detailed viewing history; coming up with a system that meets the existing SCORM standard alone may not be sufficient. Going forward, we intend to define specifications to come up with a system that meets specific on-demand video requirements.

While development efforts in the United States on the Sakai CLE continue apace, including the implementation of new frameworks, a variety of tools specific to CMS applications are being developed by institutions of higher education. We intend to continue implementing these tools in our CMS, including internationalization of them as well as localization of them into Japanese.

ACKNOWLEDGMENT

We would like to thank everyone at the Hosei University IT Research Center, the Hosei University Research Center for Computing and Multimedia Studies, and Aimluck, Inc., for their help in furthering this research.

REFERENCES

[1] Yuji Tokiwa, Yasushi Kodama, and Kazuo Yana, "Trends in open source educational system Sakai, its internationalization, and tests of scalable implementation," JeLa (Japan e-Learning Association) Journal, pp. 23-32, 2008 (In Japanese).
[2] Sakai Project, http://sakaiproject.org/, 2008.
[3] Yasushi Kodama, Tadashi Komori "An Aspect-oriented Development in Sakai" Proc. of the International Symposium on Recent Trends in Global E-learning and Collaboration, pp. 15-16, 2007.
[4] Mik Kersten and Gail C. Murphy: Atlas:a case study in building a web-based learning environment using aspect-oriented programming, Proc. of the 14th ACM SIGPLAN conference on OOPSLA'99, ACM, 1999.

Address of the corresponding author:

Yasushi KODAMA
IT Research Center, Hosei University
2-17-1 Fujimi, Chiyoda, Tokyo 102-8160
Tel: +81-(0)3-3264-4433
Fax: +81-(0)3-3264-4433
Email: yass@hosei.ac.jp

A 3-D Real-time Interactive Web-cast Environment for E-Collaboration in Academia and Education

*Billy Pham, *Ivan Ho, **Yoshiyuki Hino, **Yasushi Kodama
Hisato Kobayashi, and Kazuo Yana

**Hosei University Research Institute, California*
Burlingame, CA 94010 USA

***Hosei University IT Research Center, Tokyo 102-8160, Japan*

Correspondence: bpham@huric.org

Abstract

This paper introduces the collaborative research on Second Life, a 3-D real-time interactive web-cast environment, that has recently been undertaken between Hosei University and Hosei Research Institute, California. Through years of research on real-time interactive web-cast, our system has, since 2004, supported symposia around the world: Jetro-Biotech Japan and Jetro-(re) Structuring Success in 2003, DMS2203, The Ninth International Conference on Distributed Multimedia Systems in 2003, Ro-Man, IEEE International Symposium on Robots and Human Interactive Communications. We have established media servers in several regions, California, Paris, and Tokyo, to ensure our services for many audiences. Recently Second Life has become the best 3-D social network environment where people can do everything they wish. Following the trend and continuing our research, we have purchased two islands and set up our web-cast environment. From there, our audiences can participate more freely, and more realistically than the conventional web-cast.

1. Introduction

The rapid worldwide expansion of broadband Internet has made online communication an essential component in daily life. Students don't have to go to school to attend their classes, and businessmen don't have to travel to attend their company's meetings. Even family members don't have to pay for long distance when they want to chat with each other. The Internet really changes our world in communication. Since 2002, the Hosei University Research Institute, California (HURIC) has taken advantage of this by delivering, over the Internet, graduate-level MBA classes, in collaboration with California State University, Hayward (CSUH), from its multimedia room at Burlingame, CA, USA, to Hosei University in Tokyo, Japan. At the time, the system used a set of standard H.323 video conferencing systems and a PowerPoint synchronization system called Star Board, developed by Hitachi Software Company. However, the system was very expensive, not anytime and anywhere, and limited in its connections. Therefore, we developed a web-cast system that uses a small-delay video/audio streaming system and Java-based teaching material synchronizing/sharing application. We called the system a 2-D conventional web-cast in which our participants communicate by video and text-chat. We have used this system for years. But recently we wanted our audiences to have more realistic communication in a web-cast environment where our audiences can express their emotions, share their knowledge, and engage in friendly collaborations. Second Life has recently become more and more popular as a 3-D interactive social network environment, so we have put our efforts into researching its possibilities. Currently, we have developed a new environment in which we have user-friendly 3-D interactive communication, and realistic expression. Our new web-cast environment has been built on our Second Life islands, which are called HOSEI UNIVERISTY CAMPUS, and HOSEI RESEARCH CENTER.

2. System features

With users' imaginations, and Second Life's friendly UI, the environment has been created wildly. A newbie visiting any place in the Second Life world will be surprised by the diversity, complexity, and realistic nature of those environments. Moreover, in the world of Second Life, a user is represented by his or her avatar which is freely customized to pursue freedom, dreams or interests. Therefore, Second Life represents a new virtual life, and people will find their interest in the ultimate of exploration in the Second Life world.

In Hosei University's research, we have focused on Second Life's system features to achieve our goal.

2.1 Avatar

In Second Life, an avatar represents a user. A user can freely customize his or her avatar's look from head to toe. The user can buy or create clothes, jewelry or shoes. Most avatars are not identical so that they can be easily recognized by friends or colleagues. With a set of gestures, an avatar can perform most movements or express lifelike emotions. Amazingly, he or she can dance in a ballroom, play a piano in a concert, raise a hand to ask a question, and much more. In a web-cast conference, an avatar can represent a presenter who is standing on a podium to talk, using a mouse to change his or her PowerPoint slide, or even walk around to see the audiences' interests. On the other hand, the audience avatars can respond to the lecture by shaking their heads, laughing, or even falling asleep. Whenever they want to ask questions, they just simply raise their hands. Avatar movement is very important for a web-cast environment because it shows a connection between the presenter and the audience. The connection will enhance the effectiveness of content delivery.

2.2 Video Streaming

Video streaming in Second Life is simple. While web-casting, some presenters want to share some video files or live event video streaming. Currently, Second Life supports only QuickTime files. Therefore, live video streaming in Second Life comes from a QuickTime encoder to the QuickTime Media Server, and to the Second Life web-cast venue. For on-demand video streaming, the file comes from the QuickTime Media Server to the Second Life web-cast venue. There is only a video streaming per parcel in the Second Life environment. Video Streaming in Second Life is also very important because it will enhance the web-cast and will fulfill the event.

2.3 Content Sharing

Today, in real life, when presenting an academic paper, PowerPoint slides are essential. At Hosei University, we have developed a content sharing object which can share the PowerPoint slide as an image. The object is created by Second Life primary object and Linden Scripting Language (LSL). While web-casting, a presenter can control his or her PowerPoint slides forward or backward, or even select any slide. Moreover, we also care about authorization for the object; so the avatars' names are set in the object. This way, we can allow presenters or assistants to control the object while the audience has no control over it.

2.4 Flexible web-cast environment

Anyone can create his or her own environment in Second Life, from home to grocery store, from vacation beach to water resort. So, it is possible to create multiple web-cast environments. Changing the web-cast environment enhances audience impression. In Sky Hall, which is built on the top floor of Boissonade Tower, a symbol of Hosei University, we create our web-cast environment with a video streaming screen, PowerPoint sharing projector, lectern, chair, sofa set for break-time chatting, coffee table, painting, and much more. We also allow presenters to add images, or decorate his or her own session. He or she can create objects in 3-D to demonstrate to the audiences. In Second Life, when creating environments, people have to use some primary objects, prim. In addition, people can group up to 255 objects as an environment set, place the set in their inventory folder, and reuse it as they please. Therefore, the Second Life web-cast environment is flexible and reusable.

2.5 Multilingual support

Nowadays, the web-cast can become an increasingly international event, and there are some audiences who want to understand the content in their own language. Our environment also supports the multilingual feature. We have created an object and, using LSL script, can connect with an external translation engine such as Google. A text will be sent from a Second Life venue to the Google translation engine. From there, the engine will return the translated text. When the text is received, the object will display in the user's configuration. Our audience just touches the multilingual object, and then selects the translated language, and translated contents will appear.

2.6 Multiple channels for communication

Both voice chat and text chat are available in Second Life. In addition, Second Life allows an avatar to communicate with a group or an individual. This feature is useful because, in a web-cast, the presenter wants to talk to all members of the audience. On the other hand, an audience member sometimes wants to talk with his or her friend and doesn't want to bother anyone else. For instance, if I come in late for a web-cast session and want to catch up what is going on, I just whisper to my friend to get some information. This feature in Second Life is even better than in real life because the targeted group or individual will receive the communicated signals, and the others will not.

2.7 Full admission control

The Second Life world is truly a real world; so people do whatever they want. Sometimes people behave in unsuitable ways for an academic event. However, Linden Lab has created regulations which ensure that all avatars communicate "legally", and it allows for the virtual land owner to apply his or her own regulations. Therefore, when an avatar enters Second Life and behaves wrongly, he or she automatically creates his or her own restrictions. So when setting up a Second Life web-cast environment, an administrator can configure it. It can be opened to the public or just some invited guests. Whenever an audience member does something wrong, the administrator can ban him or her right away. However, we have never encountered such problems; all of our guests are polite and elegant.

3. Advantages

3.1 Social Network

Social networking has become a component of daily life. People connect to people virtually, look for an old friend, or make friends with someone at the other side of the world. Many companies have offered these services, with Orkut from Google, Facebook, and Hi5 among them. Nowadays, Second Life is really a social network where people can make or find their friends, where people can share their vision or learn each other's, or where people can find a place to express themselves. Moreover, Second Life offers real-time interactive features so that users find it easier to communicate with each other. Today it is becoming more and more popular, and an increasing number of educators, artisans and business people know about it. Most importantly, Second Life is a place with more than one and half million users.

3.2 Realistic Interaction and Environment

Gestures in Second Life have been created by replicating daily life movements. Walking, dancing, laughing, and much more, have been applied in default settings for each avatar. However, people can create more gestures, such as performing music and playing sports. Beside the realistic interactions, Second Life has also created realistic environments which can perform like natural places, where the environment changes by time and weather. With light and reflection effects, Second Life objects perform more realistically, based on users' imaginations. This is one of the special features of Second Life because it easily connects with users.

3.3 Multiple communication channels support everyone including people with disabilities

Second Life has a variety of communication channels. Text-chat or voice-chat can be used as a group or individually. The UI of those is so simple that anyone can use them, including people with disabilities. For people with hearing impairment, using text-chat is an appropriate solution. The people can apply colors and text-sizes into text-chat to enhance their communication. In addition, event-transcript is also valuable. The event-transcript captures lectures, or speeches and converts them to text files. For people with visual impairment, voice-chat is the best solution. Just simply connect with a web-cam, with a built-in microphone, and they can chat with everyone in the Second Life world. For people who are mobility-impaired, research in Japan has shown that they can use their brain waves to easily stroll in Second Life. Nowadays, some commercial companies have introduced products which support those people who are mobility-impaired. One of the products is HandsFree3D, which allows someone to operate in Second Life without using their hands.

4. System in practice

Prof. Hisato Kobayashi initiated our Second Life project in 2007. We quickly gathered our members and started to explore it. After that, he also engaged us in a practical event for the project, RO-MAN2007, 16th IEEE International Symposium on Robot & Human Interactive Communication, which was held in Jeju Island, Korea in August of 2007.

4.1 RO-MAN2007

This was the first 3-D web-cast in Second Life. We rented a small virtual land and set up some basic components for the web-cast, including a video streaming screen which showed the audience of the event keynote speakers. If the audience had questions or comments, they could use text-chat. At the time, the web-cast environment was simple, with some chairs placed on a grass field, and with some posters announcing the event schedule. Everything went through beautifully, and that achievement encouraged us to continue the project.

4.2 JeLA2008

After the success of the first event, we purchased two of our own Second Life islands. On the HOSEI RESEARCH CENTER island, we have replicated our institute (HURIC) in California, and HITEC building in Tokyo. On the other island, we built the Boissonade Tower, which represents our main campus in Tokyo. In the tower building, we created a Sky Hall where we planned to host our future web-cast events. In the hall, we set up a screen for live or demand video streaming, a sharing content screen, podium for the presenter, chairs for the audience, sofa sets, coffee table, hot coffee cups for break time, and much more.

JeLA2008 was the first practical event for the new environment. We operated it, synchronizing with the real event being held in Tokyo. In the event, we set up a monitor which showed the activities in Second Life so that the audiences both in Second Life and in Tokyo could participate with each other.

Figure 1. Snapshot of our Second Life islands

4.3 RO-MAN2008

RO-MAN2008 was truly our forward step. We hired a commercial company to replicate the look of our keynote speakers. We handed them pictures, and they created those avatars for us. At the time, the real conference and virtual conference were synchronized. All participants could talk to each other by voice chat. While the real presenter talked, we operated his movements by controlling his avatar.

Currently, we have more events from those organizers.

Figure 2. Snapshot of one keynote speech during Ro-Man2008 in our virtual Sky Hall on Hosei University Campus island

Figure 3. Picture of Prof. and Dr. Kazuhiro Kosuge and his avatar , a keynote speaker for RO-MAN2008

5. The events activities

While web-casting, we gave our best support to our audiences to ensure the quality of the events. We also carefully paid attention to the connections between the audiences and presenters. The audiences were quite active. They clapped their hands, shook their heads for disagreement, or raised their hands to participate. The audiences also asked many questions of each presenter at the end of his or her session. This is a confirmation of the system advantage over the 2-D web-cast system which does not have features to show realistic movements.

6. Future works and conclusion

We are happy to serve our virtual audiences, and their comments drive our future improvements.

Enhancement of speaker movement is a challenge. In order to fully capture all movements, we have to use a 3-D camcorder which is fully integrated with Second Life.

Mobility usage is also our consideration. We want to send audio and content of our events to mobile devices like cellular phones or PDAs.

Most of all, I would like to thank all our project members who have put forth tremendous effort toward our achievement, and all our audiences who always share their visions about the events.

References

[1] B. Dalgarno, “Spatial Learning in Virtual Environments”, VDM Verlag Dr. Mueller e.K., August 27th, 2008.
[2] T. Boellstorff, "Comming of Age in Second Life", Princeton University Press, April 21th, 2008.
[3] J. Heaton, "Introduction to Linden Scripting Language for Second Life", Heaton Research, Inc., December 17th, 2007
[4] J. Heaton, "Scripting Recipes for Second Life", Heaton Research, Inc., July 8th, 2007
[5] J. Heaton, "Introduction to Textures, Animation Audio and Sculpting in Second Life", Heaton Research, Inc., September 24th, 2007.

Applying Flow Theory to the Evaluation of the Quality of Experience in a Summer School Program Involving E-interaction

Kiyoshi Asakawa[1, 2] and Kazuo Yana[3]
[1]Department of Intercultural Communication, Hosei University, Tokyo, Japan
[2]Hosei University Research Institute, California, CA, USA
[3]Hosei University IT Research Center, Hosei University, Tokyo, Japan

Abstract — In 2008, we, Hosei University Information Technology Research Center (HITEC) and Hosei University Research Institute, California (HURIC), started a new project to evaluate our overseas intensive English and ICT learning program for high school students. This evaluation focuses on participants' quality of experience while attending the program. The approach has been developed from the perspective of a psychological theory, called "flow theory." This paper introduces our new project as a preliminary example of how flow theory can be applied to evaluate educational programs, what kinds of analyses can be conducted for its purpose, and what can be learned from the results for future modification and development of such educational programs toward their effectiveness and enjoyability.

I. INTRODUCTION

A. *Overview of the "California Summer English & ICT Training Program (EICT)"*

Since 2004, Hosei University Information Technology Research Center (HITEC) and Hosei University Research Institute, California (HURIC) have been closely collaborating to organize an oversea summer intensive learning program, called "California Summer English & ICT Training Program (EICT)," for students from 3 affiliated high schools of Hosei University. The EICT program is a Hosei University-High School Joint Project, and it is designed to provide effective English and ICT learning environments for the affiliated high school students. The concept of the program is to "learn ICT in English and learn English through ICT," and the program consists of two periods of sessions: (1) pre-training classes offered in Japan and (2) California summer intensive classes. For the pre-training session in Japan, participating students are provided laptop computers with English OS at the beginning of the academic year and they are given a series of classes and lectures on computer literacy in English using environments, such as communication via email, creation of home page, power point presentation, etc. They also learn practical basics of computers, such as designing and producing adders using TTL (hardware training) and JAVA programming (software training), which are normally taught at the first and second year levels at specialized math and science universities. In addition, the participating high school students interchange with American high school students who are learning Japanese at Mills High School, an affiliated high school in Millbrae, California, through videoconferencing and email during the pre-training period. In summer, the students go and stay at California State University, East Bay (CSUEB) for 3 weeks, and receive intensive ESL and High Tech. lectures by instructors of the university. During the stay in California, the high school students also deepen their friendships with Mills High School students in person through a home stay program and several occasions for interchange. For the EICT program, each student has his or her own research topic, develops it through the program, and gives a distance presentation on the topic on the last day of the training in California to their school teachers, instructors, and family members in Japan via a videoconferencing system connecting Hosei IT Center and a distance lecture room at CSUEB. All the activities and lectures of the EICT program are geared toward this final presentation, and students must carry out their own tasks independently. Since the EICT program was launched in 2004, English language skills of participating students have been evaluated yearly before and after the program and statistically meaningful improvement has been witnessed in spite of its short period of training [1].

B. *Flow Experience and its Theory*

In 2008, after 4 years of successful operation of the EICT program, we, HITEC and HURIC, have now started a new project to evaluate the EICT program, focusing on participants' quality of subjective experience from the perspective of flow theory, as a next step for its future modification and refinement. Flow theory provides a model of human development which explains how individuals develop higher levels of skills or abilities through daily experiences called "flow." Flow is an optimal state of experience in which an individual feels cognitively efficient, deeply involved, and highly motivated with a high level of enjoyment [2][3]. In this state, an individual is fully functioning by stretching existing skills while tackling optimal challenges perceived in an activity. There, most of the person's attention is directed to the activity at hand to express and realize personal potentials.

According to the theory, two criteria must be satisfied in order to experience flow [3][4][5]. That is, (1) the individual's perceived challenges of activities at hand and his or her perceived abilities or skills to take up the challenges must be in balance, and (2) the levels of the two experiential dimensions must be relatively high. In other words, the quality of subjective experience is a function of the perceived challenges and skills. Moreover, previous flow research has suggested that the flow experience has significant potential for fostering important aspects of personality, such as self-esteem and psychological resilience, developing academic and social skills and commitment, cultivating and promoting meaningful life challenges and social integration, improving the quality of life and life satisfaction, and promoting positive attitudes toward the future [6]. Therefore, according to these previous findings, it may be possible that experiencing flow in the activities and classes of the EICT program not only develops the English language skills and ICT related knowledge of participating students, but also promotes their psychological well-being if they can experience a great deal of flow while attending the program.

C. Overview of Analyses to Evaluate the EICT Program in 2008

The final goal of our project is to make the EICT program enjoyable and effective by producing the flow experience in participating students. As a first step to achieve the goal, we have started with the examination of how high school students who participated in EICT2008 (California Summer English & ICT Training Program in 2008) experienced the program. For the examination, we focused on their quality of experience, especially their flow experience. As mentioned earlier, the quality of experience is a function of the perceived challenges and skills. Thus, we also examined the students' levels of perceived challenges and skills in different activities and classes of EICT2008 as important criteria for the evaluation. Moreover, in order to examine the relationships of the flow experience and psychological wellness of the participating students, we conducted a series of correlation analyses between their average intensity of flow experience during the intensive training period in California and several indicators of psychological well-being at the end of the program.

II. METHODS

A. Participants

Twenty-one students from 3 high schools affiliated to Hosei University registered to participate in EICT2008, starting in January 2007. Among them, 7 students (4 males and 3 females) were from Hosei University High School, 6 (all males) from Hosei University Second High School, and 8 (all females) from Hosei University Girls' High School. During the training period in Japan, a student (male, senior) from Hosei University Second High School dropped out the program and a student (female, senior) from Hosei University Girls' High School joined it instead. The final participants who completed EICT2008 were a total of 21 high school students, consisting of 9 males and 12 females, of whom 19 were juniors and 2 were seniors.

B. Procedure

In order to evaluate what experiences the participating high school students had in EICT2008, we used a modified Experience Sampling Method (ESM) [7][8]. For the general ESM procedure, participants carried pre-programmed wristwatches which signaled them 8 times daily for a weekly total of 56 signals to fill out an ESF (Experience Sampling Form). The ESF was designed to elicit information on the participants' locations, activities, thoughts, and accompanying psychological states. For this study, however, we asked participants to fill out the ESFs only after the target activities and classes of EICT2008 to report their subjective experiences there: 6 times in Japan (after the pre-training classes) and 9 times during the stay in California (after 3 different intensive English classes, 3 different intensive High Tech. classes, a field trip, a preparation class for the final presentation, and the final presentation). In addition, in order to examine the students' psychological wellness, we administered a questionnaire which included measures of self-efficacy, will for meaningful life, anxiety, self-esteem, and Jujitsu-kan in daily life, at the beginning of EICT2008 (January, 2007), a week before leaving for California (July, 2008), on the last day of California intensive training (August, 2008), and 3 weeks after returning to Japan (September, 2008).

C. Measures

1) *Quality of Experience*: We examined the quality of experience with eleven experiential items on the ESF: concentration, enjoyment, interest, Jujitsu-kan (a Japanese sense of fulfillment), happiness, activeness, satisfaction, strength, excitement, perceived importance to oneself, and perceived importance to future goals. We measured the experiential items on 3 different rating scales; 4 experiential variables – happiness, activeness, strength, and excitement – were measured by 7-point semantic differential items: happy-sad, active-passive, weak-strong, and excited-bored; 3 experiential variables – concentration ("How well were you concentrating in the class?"), enjoyment ("Were you enjoying the class?"), and satisfaction ("Were you satisfied with yourself?") – by a 10 point rating scale ranging from "not at all" to "very much"; and 3 experiential variables – interest ("Was the class interesting?"), perceived importance to oneself ("How important was the class to you?"), and perceived importance to future goals ("How important was the class to your future goals?") – by a 9 point rating scale ranging from "not at all" to "very much." In addition, we measured the participating students' Jujitsu-

kan ("What level of Jujitsu-kan were you getting in the class?), perceived challenges ("Level of challenges of the class"), and perceived skills ("Level of your skills for the class") by a 9 point rating scale ranging from "low" to "high."

2) *Intensity of Flow*: In order to measure the intensity of the flow experience, we used the composite of a set of ESF experiential variables, concentration, enjoyment, interest, and Jujitsu-kan, which are expected to be concurrently high in flow [9]. Interest and Jujitsu-kan were measured on a 9 point rating scale (1 – 9) and concentration and enjoyment were measured on a 10 point rating scale (0 – 9). Thus, we first standardized raw scores of each variable by using sample means of the variables, and then summed up the resulting standardized scores of these variables to make a continuous measure of flow. We labeled this composite variable as "flow-score," and the higher the flow-score is, the more intensely a participant experiences flow.

3) *Self-efficacy, Will for Meaningful Life, Anxiety, Self-esteem, and Jujitsu-kan in Daily Life*: In order to examine psychological wellness of the participating students, we administered a battery which included measures of self-efficacy, will for meaningful life, anxiety, self-esteem, and Jujitsu-kan, several times over the entire EICT2008 program. We used a Japanese version of Sherer's self-efficacy scale [10], translated by Narita et al. [11], to measure their self-efficacy, used Kondo and Kamata's will for meaningful life scale [12] to measure their will for meaningful life, used a Japanese version of Spielberger's State Trait Anxiety Inventory [13], translated by Shimizu and Imasakae [14], to measure their tendency to generate anxiety in daily life, and used a Japanese version of Rosenberg's self-esteem scale [15], translated by Yamamoto, Matsui, and Yamanari, [16], to measure their overall self-esteem. In addition, we used a single item, "I think my life is fulfilling (or I am getting Jujitsu-kan in daily life)," to measure the students' Jujitsu-kan, a Japanese sense of fulfillment, in their daily lives [6].

III. RESULTS and DISCUSSION

A. *Quality of Experience in Different Classes and Activities of EICT2008*

The main purpose of our project reported in this paper was to evaluate a Hosei University-High School Joint Project, "California Summer English & ICT Training Program in 2008," by examining participants' quality of experience from the perspective of flow theory. For the purpose, first the high school students' quality of experience was compared between 5 categories of activities in EICT2008; (1) class activities in Japan, (2) class activities in California, (3) a field trip in California, (4) a preparation class for the final presentation, and (5) the final presentation. Table 1 shows the results. The high school students' levels of perceived challenges and skills were the highest in the final presentation, whereas their perception of challenges was the lowest in the field trip and their perception of skills was the lowest in the classes taught in Japan (challenges, $F(4, 307) = 5.52$, $p < .001$; skills, $F(4,307) = 2.57$, $p < .05$). The levels of concentration, enjoyment, Jujitsu-kan, and flow experience as the composite of concentration, enjoyment, interest, and Jujitsu-kan were also the highest in the final presentation, whereas the levels of concentration and Jujitsu-kan were the lowest in the classes in Japan and the levels of flow and enjoyment were the lowest in the preparation class for the final presentation (concentration, $F(4, 307) = 4.01$, $p < .01$; enjoyment, $F(4, 306) = 3.64$, $p < .001$; Jujitsu-kan, $F(4, 307) = 3.92$, $p < .001$; flow, $F(4, 304) = 2.66$, $p < .05$). As for the levels of happiness, activeness, strength, excitement, perceived importance to themselves, and perceived importance to future goals, the high school students scored the highest in the final presentation, while they scored the lowest in the preparation class for the final presentation for happiness and activeness, scored the lowest in the classes taught in Japan for strength and excitement, and scored the lowest in the field trip for perceived importance to themselves as well as to their future goals (happiness, $F(4, 307) = 6.44$, $p < .001$; activeness, $F(4, 307) = 5.10$, $p < .001$; strength, $F(4, 307) = 6.74$, $p < .001$; importance to oneself, $F(4, 306) = 3.30$, $p < .05$; importance to future goals, $F(4, 306) = 4.01$, $p < .01$). According to these results, the students appeared to have the most positive experiences in the final presentation where their perceptions of challenges and skills were both the highest. They showed higher concentration, felt more enjoyment, felt more Jujitsu-kan, felt happier, more active, stronger, and excited, felt more personal relevance and to their future goals, and most importantly they experienced flow more intensely in the final presentation, as compared to other types of activities. On the other hand, the high school students appeared to have relatively fewer positive experiences in the classes in Japan and the preparation class for the final presentation as compared to other activities of EICT2008.

Table 1 High school students' quality of experience in different activities of EICT2008

	Classes in Japan	Classes in California	Field Trip	Preparation for Presentation	Final Presentation	*F*
	(N = 120)	(N = 125)	(N = 21)	(N = 21)	(N = 21)	
Challenges	6.29	6.63	5.43	7.05	7.86	5.52***
Skills	4.76	5.20	5.52	5.24	5.90	2.57*
(Challenge – Skill)	1.53	1.43	-.10	1.81	1.95	1.90
Flow (composite)	-.37	.13	.16	-.88	1.88	2.66*
Concentration	6.60	7.14	7.00	7.33	8.00	4.01**
Enjoyment	7.11	7.13	7.48	5.71	7.67	3.64**
Interest	6.78	6.65	6.33	6.24	7.38	1.36
Jujitsu-kan	6.32	6.85	6.95	6.52	7.76	3.92**
Happiness	4.98	5.21	5.38	4.10	6.05	6.44***
Activeness	4.65	5.13	5.24	4.64	5.95	5.10***
Satisfaction	5.77	6.17	6.43	6.24	6.86	1.63
Strength	4.49	5.16	5.19	4.52	5.71	6.74***
Excitement	4.87	5.35	5.57	5.00	6.00	4.19**
Importance to oneself	7.02	7.02	6.71	7.76	8.19	3.30*
Importance to future goals	6.50	5.98	5.00	6.67	7.10	4.01**

Reported p-levels are two-tailed. * $p < .05$; ** $p < .01$; *** $p < .001$.

B. Relationships between High School Students' perceptions of Challenges and Skills and the Overall Quality of Experience

Flow theory postulates that the quality of subjective experience increases as perceived challenges and skills increase in balance. Thus, in order to examine how the quality of experience changed as perceived challenges and skills changed, a series of correlation analyses was performed between the high school students' perceived challenges and skills and their overall quality of experience in EICT2008. As Table 2 shows, the students' perceived skills were positively associated with the intensity of flow (r = .42, p < .001), concentration (r = .30, p < .001), enjoyment (r = .28, p < .001), interest (r = .30, p < .001), Jujitsu-kan (r = .44, p < .001), happiness (r = .21, p < .001), activeness (r = 41, p < .001), satisfaction (r = .51, p < .001), strength (r = .43, p < .001), excitement (r = .17, p < .01), importance to oneself (r = .24, p < .001), and importance to future goals (r = .27, p < .001). That is, as their perception of skills increased, the students showed higher concentration, felt more enjoyment, showed more interest, felt more Jujitsu-kan, felt happier, more active, stronger, and excited, felt more personal relevance and to their future goals, and experienced flow more intensely. On the other hand, the students' perception of challenges was positively associated only with their perception of importance to themselves (r = .11, p < .10: marginal significance), but otherwise it was negatively associated with enjoyment (r = -.10, p < .10: marginal significance), happiness (r = -.12, p < .05), activeness (r = -.15, p < .01), satisfaction (r = -.12, p < .05), and strength (r = -.17, p < .01). In other words, as their perception of challenges increased, the students felt more importance of the activities to themselves, but they felt less enjoyment, felt less happy, less active, and less strong, and felt less satisfaction with themselves. As mentioned earlier, flow theory assumes that the quality of subjective experience increases as perceived challenges and skills increase in balance. However, the results showed that the quality of experience increased only as perceived skills increased, but not as perceived challenges increased. This is probably because the levels of challenges were quite high for the high school students throughout EICT2008. Indeed, as shown in Table 2, the average level of perceived challenges was 6.52 on the 9 point rating scale. Thus, it is quite possible that the students were almost always in situations where perceived challenges were higher than perceived skills when attending the EICT2008 classes. Flow theory labels such situations as an "anxiety" state. Then, in such situations, as the students' level of perceived skills increased and approached the level of perceived challenges, their quality of experience would have increased, while in the same situations, if their level of perceived challenges further increased, their quality of experience must have further become detrimental for the students, as the results showed.

Table 2 Correlations between high school students' levels of skills and challenges and the quality of experience while attending classes of EICT2008

Quality of experience	Skills	Challenges	*M*	SD
Skills	1.00	-.13*	5.10	1.81
Challenges	-.13*	1.00	6.52	1.91
Flow (composite)	.42***	.01	-.01	3.24
Concentration	.30***	.09	6.99	1.71
Enjoyment	.28***	-.10~	7.08	1.89
Interest	.30***	-.01	6.70	1.83
Jujitsu-kan	.44***	.06	6.69	1.74
Happiness	.21***	-.12*	5.11	1.36
Activeness	.41***	-.15**	4.98	1.42
Satisfaction	.51***	-.12*	6.08	2.10
Strength	.43***	-.17**	4.90	1.38
Excitement	.17**	-.08	5.20	1.42
Importance to oneself	.24***	.11~	7.13	1.73
Importance to future goals	.27***	.09	6.24	2.09

n = 308. Reported p-levels are two-tailed. ~ p < .10; * p < .05; ** p < .01; *** p < .001.

C. Can High School Students' Perceptions of Challenges and Skills Predict Their Flow Experience in the EICT2008 Classes?

In order to examine how the high school students' levels of perceived challenges and skills would predict their intensity of flow experience in different classes of EICT2008, a series of regression analyses was performed. Table 3 shows the standardized regression coefficients for perceived challenges and skills, predicting the intensity of flow in the classes in Japan, English classes in California, High Tech. classes in California, and the final presentation. As expected from the previous analyses, the level of perceived skills was a significant predictor of flow in all of these classes in EICT2008 (classes in Japan, *beta* = .29, p < .001; English classes in CA, *beta* = .32, p < .05; High Tech. classes in CA, *beta* = .52, p < .001; final presentation, *beta* = .54, p < .05). In other words, if a student's perception of his or her own skills was higher than that of the other students in these classes, it can be predicted that the student's intensity of the flow experience was higher than that of his or her counterparts in the classes. As for the students' level of perceived challenges, it predicted only the intensity of flow in the classes taught in Japan (*beta* = .30, p < .001). This result may indicate that the students perceived the levels of challenges, probably most of the time, in the English classes in CA, the High Tech. classes

Table 3 Regression analyses predicting high school students' flow experience by their perceived challenges and skills in different classes of EICT2008

	Standardized Coefficients			
	Classes in Japan	English Classes in California	High Tech. Classes in California	Final Presentation
Challenges	.30***	.07	-.06	.05
Skills	.29***	.32*	.52***	.54*
Adjusted R^2	.13	.06	.25	.25
F	10.11***	3.11~	10.84***	4.25*
df	118	62	59	20

Reported p-levels are two-tailed. ~ p < .10; * p < .05; ** p < .01; *** p < .001.

in CA, and the final presentation as very high as compared to their skills, but they might sometimes perceive that challenges were relatively low in the classes taught in Japan. As shown in Table 1, the average level of perceived challenges in the classes in Japan was indeed the lowest among the other activities, except for the field trip. Thus, the students might sometimes need to be more challenged in the classes taught in Japan, and as a consequence, as the level of perceived challenges increased in the classes, the intensity of flow might have increased, as flow theory postulates.

D. High School Students' Quality of Experience while Attending Intensive English and High Tech. Classes in California

EICT2008 provided the high school students with intensive English and High Tech. classes while they were in CA. The classes were the core of the program. Thus, in order to obtain a general idea of how the high school students experienced the classes, their quality of experience was examined and compared between the two classes of the California training session. Table 4 shows the results. First of all, the high school students' levels of the intensity of flow, concentration, enjoyment, interest, Jujitsu-kan, happiness, excitement, perceived importance to themselves, and perceived importance to their future goals in the English classes were significantly higher than those in the High Tech. classes (flow, $F(1, 122) = 14.20$, $p < .001$; concentration, $F(1, 124) = 14.43$, $p < .001$; enjoyment, $F(1, 123) = .10.78$, $p < .001$; interest, $F(1, 123) = 11,70$, $p < .001$; Jujitsu-kan, $F(1, 124) = 3.20$, $p < .10$ (marginal significance); happiness, $F(1,124) = 6.49$, $p < .05$; excitement, $F(1,122) = 5.36$, $p < .05$; importance to oneself, $F(1, 124) = 27.61$, $p < .001$; importance to future goals, $F(1, 124) = 29.07$, $p < .001$). That is, the high school students experienced flow more intensely, showed higher concentration, felt more enjoyment, showed more interest, felt more Jujitsu-kan, felt happier and more excited, felt more personal relevance, and felt more importance to their future goals while attending the English classes than attending the High Tech. classes. More interestingly, although there was no difference in their perceived skills between the English classes and the High Tech. classes, the students rated the level of challenges significantly higher for the High Tech. classes than for the English classes ($F(1, 124) = 10.27$, $p < .01$). Indeed, as shown in Table 4, the average level of perceived challenges in the High Tech. classes was extremely high (7.18 on a 9 point rating scale). Moreover, the difference between the levels of perceived challenges and perceived skills was significantly bigger in the High Tech. classed than in the English classes ($F(1, 124) = 7.29$, $p < .01$). According to flow theory, in order to experience flow, perceived challenges and skills must be in balance and both perceived challenges and skills must be relatively high, as discussed earlier. Thus, the relatively less positive experiences the high school students had in the High Tech. classes might be resulted from the extreme high challenges and the consequential unbalance between the levels of perceived challenges and skills in the classes. The students might be in an experiential condition called "anxiety," in a flow theory's term, when they were attending the High Tech. classes offered in California.

Table 4 High school students' quality of experience while attending English and High Tech. classes in California

Quality of experience	English ($N = 63$)	High Tech. ($N = 62$)	*F*
Challenges	6.10	7.18	10.27**
Skills	5.37	5.03	0.96
(Challenge – Skill)	0.73	2.15	7.29**
Flow (composite)	1.29	-1.09	14.20***
Concentration	7.71	6.56	14.43***
Enjoyment	7.67	6.57	10.78***
Interest	7.25	6.03	11.70***
Jujitsu-kan	7.14	6.55	3.20~
Happiness	5.52	4.89	6.49*
Activeness	5.26	5.00	0.90
Satisfaction	6.33	6.01	0.67
Strength	5.23	5.10	0.27
Excitement	5.63	5.02	5.36*
Importance to oneself	7.83	6.21	27.61***
Importance to future goals	7.00	4.95	29.07***

Reported *p*-levels are two-tailed. $\sim p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$.

E. Relationships between High School Students' Flow Experience while Attending Classes in California and their Self-efficacy, Will for Meaningful Life, Anxiety, Self-esteem, and Jujitsu-kan in Daily Life

Previous flow research has suggested that the flow experience has significant potential for helping individuals to develop psychological well-being. Thus, in the last part of the analysis, we examined the relationships between the high school students' intensity of the flow experience while attending the classes in California and several indicators of psychological well-being measured on the last day of California intensive training. Table 5 shows partial correlations between the variables, controlling for the effects of the same well-being indicators measured one week before the students left for US. As shown in the table, the high school students' flow experience in the California intensive classes was positively associated with their self-efficacy ($r = .42$, $p < .10$ (marginal significance)), will for meaningful life ($r = .48$, $p < .05$), and Jujitsu-kan ($r = .53$, p

Table 5 Partial correlations between high school students' flow experience while attending classes in California and their self-efficacy, will for meaningful life, trait-anxiety, self-esteem, and Jujitsu-kan at the end of California training

	Partial r^a	*M*	SD
Self-efficacy	.42~	72.00	15.99
Criterion	.48*	27.61	4.75
Trait-anxiety	-.72***	45.50	9.77
Self-esteem	.33	29.97	7.60
Jujitsu-kan in daily life	.53*	3.39	.78

n = 21. Reported *p*-levels are two-tailed. $\sim p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$.
[a] Partial correlations controlling for the effects of self-efficacy, will for meaningful life, trait-anxiety, self-esteem, and Jujitsu-kan, measured one week before leaving Japan.

< .05), and it was negatively associated with their trait-anxiety ($r = -.72$, $p < .001$). That is, if a high school student experienced flow more intensely than his or her counterparts while attending the California intensive classes, the student felt a higher self-efficacy, showed a stronger will for meaningful life, felt a stronger Jujitsu-kan, and felt less anxiety than his or her counterparts at the end of EICT2008. Thus, experiencing flow while attending the classes appeared to have some impact on the high school students' psychological wellness and, if this is really the case, the EICT program should be designed to promote the flow experience in participants, probably by adjusting the levels of challenges of the activities and skills to tackle them from the perspective of flow theory.

IV. CONCLUSION

Our project aimed to evaluate "California Summer English & ICT Training Program in 2008," by examining participants' quality of experience from the perspective of flow theory. The results showed that the level of challenges appeared to be very high for the participating high school students throughout EICT2008, as compared to the level of their skills, especially in the High Tech. intensive classes offered in California. In order to experience flow, the levels of perceived challenges and skills must be both high and in balance. Thus, we probably need to consider how to adjust the levels of challenges and the students' skills in the High Tech. classes. We may provide more intensive and higher level High Tech. classes in Japan for the students to prepare for the California classes, or we may lower the level of California High Tech. classes. The good news is that the students appeared to have the most positive experiences in the final presentation, even though they perceived the challenges there as the highest among all of the activities and classes in EICT2008. This means that the students had developed a relatively high level of skills through the EICT program to tackle such high challenges in the final presentation. Moreover, their overall flow experience in the California intensive training appeared to boost their self-efficacy, will for meaningful life, Jujitsu-kan, and lowered their anxiety at the end of the program. In sum, although this project has several limitations such as a small sample size and insufficient numbers of activities and classes sampled for the analyses, we strongly believe that the evaluation of the EICT program from the perspective of flow theory is a quite promising approach which provides us a great deal of information on how the participating students experience it and how they cultivate themselves through the program.

REFERENCES

[1] HITEC, & HURIC (2007). *Open research center report: Outline of research results in 2002-2006*. Hosei University.

[2] Csikszentmihalyi, M. (1975/2000). *Beyond boredom and anxiety*. San Francisco: Jossey-Bass.

[3] Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper and Row.

[4] Csikszentmihalyi, M. (1988). The flow experience and its significance for human psychology. In M. In M. Csikszentmihalyi & I. S. Csikszentmihalyi (Eds.), Optimal experience: Psychological studies of flow in consciousness (pp. 15-35). New York: Cambridge University Press.

[5] Csikszentmihalyi, M. (1997). *Finding flow: The psychology of engagement with everyday life*. New York: HarperCollins.

[6] Asakawa, K. (2009). Flow experience, culture, and well-being: How do autotelic Japanese college students feel, behave, and think in their daily lives? Journal of Happiness Studies. doi: 10.1007/s10902-008-9132-3.

[7] Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. *Journal of Nervous and Mental Diseases*, *175*, 526-536.

[8] Csikszentmihalyi, M., Larson, R., & Prescott, S. (1977). The ecology of adolescent activity and experience. *Journal of Youth and Adolescence*, *6*, 281-294.

[9] Asakawa, K., & Nakamura, J. (2008, July). *The study of autotelic personality*. Paper presented at the 4th European Conference on Positive Psychology, Rijeka, Croatia.

[10] Sherer, M., Maddux, J. E., Mercandante, B., Prentice-Dunn, S., Jacobs, B., & Rogers, R. W. (1982). The self-efficacy scale: Construction and validation. *Psychological Reports*, *51*, 663-671.

[11] Narita, K., Shimonaka, Y., Nakazato, K., Kawaai, C., Sato, S., & Osada, Y. (1995). A Japanese version of the generalized self-efficacy scale – Scale utility from the life-span perspective -. *Japanese Journal of Educational Psychology*, *43*, 306-314 (in Japanese).

[12] Kondo, T., & Kamata, J. (1998). The sense of a life worth-living among contemporary college students and its scale. *Japanese Journal of Health Psychology*, *11*, 73–82 (in Japanese with English abstract).

[13] Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the state-trait anxiety inventory(Self-evaluation questionnaire)*. Palo Alto, California: Consulting Psychologists Press.

[14] Shimizu, H., & Imasakae, K. (1981). Development of the Japanese edition of the Spielberger state-trait anxiety inventory (STAI) for student use. *The Japanese Journal of Educational Psychology*, *29*, 348–353 (in Japanese).

[15] Rosenberg, M. (1965). *Society and the adolescent self-image*. New Jersey: Princeton University Press.

[16] Yamamoto, M., Matsui, Y., & Yamanari, Y. (1982). The structure of perceived aspects of self. *The Japanese Journal of Educational Psychology*, *30*, 64–68 (in Japanese).

Proceedings
International Workshop on Visual Languages Computing (VLC 2009)

Co-Editor

Gem Stapleton, *University of Brighton, UK*

Extracting Hot Events from News Feeds, Visualization, and Insights

Zhen Huang and Alfonso F. Cardenas, University of California, Los Angeles

***Abstract*— Our aim is a real time news extraction system capable of identifying up-to-date "hot news" from large amounts of news reports on the internet. We show two important characteristics of hotness: short-term burst and long-term historical variation. Based on these factors, we define our problem: given all news articles on the current day *p* and previous news archives, identify the hot news on day *p*. Our novel system consists of four parts. First, all entities (atomic elements in text such as the names of persons, organizations, locations, etc) which occurred on the current day are extracted. Second, hot terms on day *p* are identified with our novel term weighting scheme for hotness. Third, each news article on the current day will be represented as a vector of hot terms with frequency. The last step is to cluster all new articles on day *p* and rank the result clusters with their hotness score. The scheme is also adaptive since the hot terms extracted are different from day to day, and reflect the hot events on each day. The focus of this paper is to illustrate the visualization and the insight of the hot-news through time, not the underlying algorithm details we developed.**

Our experiments conducted on a Yahoo! News dataset indicate that our term weighting scheme incorporating the two important features of hotness evaluate the significance of each extracted entity properly.

***Index Terms*— Information retrieval, visualization**

I. INTRODUCTION[A1]

There are different kinds of news from our daily newspaper and news agents. A large portion of news articles are about the topic that constantly appear, almost every day, like Iraq and Congress. Some of the news articles were only interesting to a small amount of people. They are there just because of the integrity of news. There are much fewer reports about it.

We observe that there may be a burst of news articles about an event that rarely or never has been mentioned before. Furthermore, an event cannot be hot forever. Some hot news might be hot for only one day and submerged in the sea of information the next day, while some others might last for several days. Some hot topics never disappear and become daily topics since their appearance, like "*iphone*" and "*stimulus plan of the current economy*". Hot events which last for several days would be considered hot on the first few days they appear and less hot the following days with the decay of their hotness and emergence of other hot events. In a streaming environment, the burst size and different stages of burst need to be considered for hotness.

Fig. 1 gives the definitions of terms we use. The hotness of a certain event will decay along with the elapse of time. An article can be estimated by a set of certain entities shown in the article. In order to examine the burst of events, we estimate it by examining the burst of certain terms. The burst of a term is related to the number of documents in which it occurs and the time length of being popular. The historical variation of a term indicates how often this term appeared in history. By combining these two factors together, we can compute the hotness scores for all entities which occur on the current day and get a list of hot terms for that day. With the hot terms, we cluster the news articles which are featured by these hot terms and rank the result clusters by their hotness score.

Hotness definition

There are many definitions of the concept "hotness" [3] [11] [12]. One of the intuitive ideas is that hot events are the events on which a lot of news articles are written and thus we only need to cluster the news articles on the current day and choose the top clusters in size as the hot events. However, it is not necessarily the case that the larger a cluster is, the hotter that subject will be. We have to take the short-term burst and the historical variation into consideration. For example, there were 24 news articles about Iraq on Jan 04, 2007, which is the largest cluster that day. However, it was not the hottest subject because there was a lot of news about Iraq in history constantly.

We define "hotness": a burst of the number of articles related to an event during a short time span, wherein the event did not occur often (historical variation) in history. As discussed in Allan et al. [4], event is defined as "some unique thing that happens at some point in time". News articles about an event could be anything related to this event. It could be news reporting this event from different sides. There is a clear distinction between an event and a topic. As shown in Fig. 1, topic is defined as a broader category of events, or a class of events. "Volcanic eruption" is a topic, while "the eruption of Mount Pinatubo on June 15th, 1991" is an event [4].

Document clustering

There are many methods proposed for document clustering [5]. HAC (Hierarchical Agglomerative Clustering) [5] is one

Alfonso F. Cardenas, Author, Computer Science Department, University of California, Los Angeles, Ca 90024 (email: cardenas@cs.ucla.edu)

Author Zhen Huang, student Computer Science Department, University of California, Los Angeles, Ca 90024 (email: zhenhuang@ucla.edu)

of the hierarchical methods and starts with each article being placed into its own cluster and greedily chooses the closest pair of clusters to merge one at a time. To select the closest pair of clusters, there are many similarity comparison methods, such as Dice, Jaccard, Overlap, Cosine, and L1 Norm. Although HAC can model arbitrary shapes and different sizes of clusters, it cannot be used directly for news stream applications due to the following two reasons [1]. First, HAC builds a dendrogram which then relies on human beings to partition it to produce actual clusters. This step is usually done by human visual inspection, which is a time-consuming and subjective process. Second, the computational complexity of HAC is expensive since pair-wise similarities between clusters need to be computed.

Due to the frequent update and streaming nature of a news source, all text mining methods on news streams should be updatable. Much research [2] [3] has been focused on identifying hot news through exploring news articles from different news sources. Their case is different from our case because many other characteristics could be used by them, such as news source reputation. The data set we use is from the Yahoo! News RSS feed. It serves as an aggregator of the news from different sources. All the news articles are mixed together and their origins are unknown.

Article representation

There are several ways to represent a document. For example, a document could be represented as a bag of words which appear in the article or a vector of distinct terms with their weight (significance). The sequence of words and the structure of the text such as sentences and graphs are ignored. Because of different purposes, different words are extracted from the source article. For *tf-idf (term frequency-inversed document frequency)* which is used in the document similarity measurement, all terms that appear in the article need to be considered. In the news environment, entities such as persons, organizations, and places play an important role in identifying an event or news article in our case. "Hot" terms are entities that are contained in news articles about a hot event. To weigh how hot an event is, we have to weigh how hot the entities about the event in the news article are. There are many ways to evaluate the significance of a term. A detailed description of our term weighting scheme is presented in Section 4.

II. System Overview

The goal of our system is this: given all news articles on the current day p and previous news archives, identify the hot news on day p. Fig. 2 shows the overview of our system. The raw data is all news articles on the current day p. After preprocessing, each news article is represented by a vector of words with their term frequency in this article. The preprocessing includes stop word removal and named entity extraction. With the entities extracted from each article on day p and their historical occurrence extracted from the historical news article achieve, we identify the hot terms on day p with our novel term weighting scheme for hotness. The historical occurrence of the entities is needed here because they are used to compute the stage of the burst and the historical variation, which are the two key characteristics of hot terms by our hotness definition.

Since a large portion of terms are not hot terms, only terms with a hotness score above a certain threshold are extracted and the results are put in a list called the "traceable list". With the "traceable list" and their hotness score, we represent each news article on day p by a vector of hot terms which occurs in that articles and is in the "traceable list". Since only the terms in the traceable list are considered, the dimensions of the feature vector of each news article are greatly reduced and the consumed computation resource is highly reduced.

With each news article on day p represented by its feature vector, the hierarchical agglomerative clustering (HAC) method is employed to group the articles into clusters. Since each cluster contains all news articles about an event and related events, it is also given a hotness score to show how hot this event is. The highest hotness score of the terms in the news articles in the cluster is considered as the hotness score of the cluster. With the hotness score given to each cluster, we can rank all clusters with the score from highest to lowest. At this stage, we can answer the user's question: "what is the hot news today?"

Furthermore, our system works in a real-time fashion with one day as the basic time unit. The hot terms and hot new articles extracted are different from day to day, and reflect the hot events on each day.

III. Term Weighting Scheme

After preprocessing in our system, each news article is represented by a vector of entities with their term frequency. The next step of our system is to derive a brand new term weighting scheme for the hot term identification from all terms which occur on the current day. There are many term weighting schemes proposed for different purposes, such as *TF-IDF* [6], *TF-ISF* [7], *TF-PDF* [8], etc.

We introduce a new term weighing approach with the following heuristics:

- The method should be *updateable*. Hundreds of news is published by single news agents every day. With the growth of news articles, it is essential to develop incremental methods to extract hot events or hot topics. The dataset is not divided into a training set and a test set. The historical statistics should be updated adaptively. A hot event does not remain hot forever, but only during certain time periods.
- Hot event should be detected early and be adaptable to the life cycle of an event
- Parameters do not need to be manually set
- Consuming minimal computation resources

Instead of consuming a lot of computation resources like Topic Similarity Measurement and document clustering, we provide an efficient and fast-adaptive approach that we shall not detail in this paper. [3] is only interested in extracting "hot topics" from a given set of text-based news documents published during a given time period. Our approach can work in a streaming fashion. Furthermore, instead of identifying targeted news article as hot or not hot [3], we give it a hotness score.

We introduce our two critical properties of a hot term: "short term burst" and "long term variation (rareness)". In news streams, newly shown terms have significant importance. For instance, a unique term in a news stream may imply a new technology or event that has not been mentioned in previous articles. For each named entity, we compute the short-term score. This is considered to be the burst score. We also compute the historical variation score, which is directly related to how often this term occurred in history. These two scores are combined then to get the actual hotness score for this term.

Our approach examines all factors related to the definition of hotness and follows the following rules:

1. Important (hot) terms appear more frequently within a document than unimportant terms do.
2. From the view of a short term period, the more times a term occurs in all news articles, the stronger its discriminating power becomes. A burst of certain terms is a strong indicator of an event.
3. From the view of a long term period, the less variation a term has throughout history, the weaker its discriminating power becomes.

In this way, terms like "*Iraq*", "*U.S.*" should get a lower score. Another example is the term "*iphone*". When the term "*iphone*" was first introduced on Jan 09, 2007, it should have a high score. As time goes on, there are many news articles about "*iphone*" every day. The term "*iphone*" became more pervasive and has less variation of its occurrence, and thus it is seen as less hot.

Short-term score (burst score) computation

The number of term appearances in each document is not considered as a factor of term hotness. No matter how many times this term appears in a document, it is just in one article. The more times this term appears in one article, it is highly possible that this term is the "topical" term for this article. However, it doesn't mean this term is hot. To measure the burst of a term, it is more reasonable to consider the short-term document frequency over the days of certain window size, which is the total number of documents containing this term each day. The total number of documents each day is more important than the term frequency in one new article when judging the hotness of a term. For example, if the term "Steve Jobs" appears many times in one news article, this frequency does not make this term more hot if "Steve Jobs" only appears in that article on that day. For each entity, its short term score is computed with a formula we proposed base on this rationale:

1. The more news articles contain this term (document frequency) on that day, the hotter this term is on that day.
2. If two terms have the same score value, the tie is broken by the number of news articles containing the term on day *(p-1)*. News articles containing the term on previous days (of *window_size*) should also be taken into consideration. This is natural for an event in a news article. Some events last only one day, while some others last more than one day.
3. The more days before time point p, the less contribution it will make towards the hotness score of the same event on day p. The document frequency of term t on day p should have more weight than the document frequency of the same term on day *(p-1)*. Viewers will lose interest in the same event eventually, so the hotness may decay exponentially with the elapse of time.

The window size is set to four days based since through observation we saw that news articles related to the same event usually do not last longer than three days. We also assume that the articles on the previous four days have a positive effect on hotness, while documents further back have a negative effect on the hotness. The longer an event lasts, the less hot the event would be.

In our approach, one day is considered as a basic time unit when the document frequency is measured, which is natural for news articles. The update rate of news articles in the news streams tends to be one day in general.

Historical variation score computation

In order to measure the hotness of a term, its long term variation also needs to be considered, which is denoted as *hist_vari(t, p)*. A large variation shows that the term was more burst-related, rather than terms like "*U.S.*" which appears in our news articles every day or every now often. A term with large variation denotes that there was a burst of this term (event) in history. The variation could be the number of days a term occurred or the number of documents this term appeared in. We hypothesize that a larger variation leads to a hotter term. There are several factors which might be related to its historical variation that we examined. We introduce the following concepts, definitions, and formulations elsewhere to compute the historical variation score. Examples of the needed visualizations for these are below.

Hotness score computation

As mentioned earlier, each news article is represented by a vector of entities which contain information about people, locations, organizations, etc. First, we obtain a list of terms occurring on day p. For each term t, we obtain its short term score (burst score) and historical variation score with the above mentioned methods. The next step is to compute the hotness score of all terms that occurred on day p, formally *hotness_score(t, p)*. As discussed above, the greater the burst score, the larger the burst size, the hotter the term. Similarly, the larger historical variation the term has, the less common the term, and the hotter the term. Thus, the hotness score of term t is computed as:

$$hotness\_score(t,p) = burst\_score(t,p) \cdot historical\_variation(t,p)$$

With the hotness score of each term computed, the terms which occur on day *p* are ordered from highest to lowest by their hotness score. Terms with a hotness score above certain threshold are extracted and the list is called the "traceable list".

Our term weighting scheme is updatable and thus adaptive. The "traceable lists" are different from day to day. For a certain term, it may be hot on some day with the burst of an event, while not as hot on some other days with the decay of this burst. Our weighting scheme reflects the life cycle of a burst and the fast-paced news environment, in which fresh news emerges every day. Our scheme also reflects the historical variation of a term. More common terms like "*U.S.*", "*Iraq*", and "*California*" appear less in our "traceable list", except when there is really a burst of these terms. An example is when President Bush delivered a speech on Iraq which drew global attention on Jan 09, 2007. This case can also be modeled correctly in our scheme.

IV. NEWS ARTICLE CLUSTERING

Since the hottest terms are extracted each day, the next step is to identify the hot articles from all news articles on day *p*. During the first phase, each article on day *p* is represented as a vector of terms and their term frequency. We represent each article in a different way: a vector of terms which were in the article and also in day *p*'s traceable list. Term weight is given by the following equation:

$$term\_weight(t,d,p) = hotness\_score(t,p) \cdot normalized\_term\_frequency(t,d)$$

term_weight(t, d, p) is the term weight of term *t* in document *d* on day *p*; h*otness_score(t, p)* is the hotness score of term t on day p; *normalized_term_frequency(t, d)* is normalized term frequency of term *t* in document *p*. The term frequency of term *t* is taken into consideration because terms which occurred more frequently in an article tend to be topical terms for this article. The more a term appears in an article, the more important this term is to the whole article. Fig. 5 shows the representation of an article.

The next step is to cluster the articles. The vector is represented as a feature vector for this article for clustering.

The Hierarchical agglomerative clustering method is used to group articles into clusters of same events. The cosine similarity measure is used on as our similarity measurement. The next step is to rank the result clusters. In order to do that, each cluster has a set of terms which appear both in day *p*'s traceable list and the articles of this cluster. At first, we try to give the hotness score of a cluster as the average hotness score of all terms that appear in both the traceable list and the cluster. The highest hotness score of the term in the set is considered as the hotness score of the cluster to account for clusters that have very few but very hot terms. Formally, the hotness score of a cluster is given by:

$$\Re = \{terms\ appears\ in\ the\ traceable\ list\ and\ the\ articles\ in\ cluster\ c\}$$

$$hotness\_score(c,p) = \max\{hotness\_score(t_i,p) \mid t_i \in \Re\}$$

V. HOT EVENTS EXPERIMENTS AND VISUALIZATION

In order to evaluate the accuracy rate of the hot news articles generated, we implemented a real-time hot subject extraction and rank system in Java. The system pipeline is implemented with the IBM UIMA framework [9]. We evaluated our system against the Yahoo! data sets. LingPipe [10] is used for entities detection (e.g. people, organization and location) from each news article.

We explore the effectiveness of our approach to identify and rank "hot" entities and consequently "hot" subjects for the given date, e.g., what are the hot terms and hot subjects on a certain date, such as Jan 04, 2007. Given the date Jan 04, 2007, only the news articles on and before that day are available to us.

We first examine the set of entities generated from an article to see whether the set covers most entities shown in the article such as people, locations, and organizations involved. Then, the hotness score of each term on day p is computed. With the hotness scores, we get the hot terms on that day and cluster the news articles, which are represented by the extracted hot terms. By comparing with manually tagged hot news articles, we get the accuracy recall rate of our algorithm.

Dataset description and overview of experiments

The Yahoo! news dataset is a collection of 125,871 news articles from all Yahoo! News RSS feeds, collected between 01/2006 and 03/2007. Through LingPipe, which implements a dictionary-based chunker that performs approximate matching, 317,475 distinct entities are recognized and extracted. Other words and the order of each word are not considered. LinePipe extracts entities effectively. Most of event-related entities are extracted. Fig. 6 shows the top 50 entities with highest term appearance in our dataset. We filtered out 274,192 infrequently occurring entities by requiring that an entity should occur twice in an article, leaving 43,284 entities in the dataset.

In order to fully validate our system, we conducted three experiments. First, we compared hot terms extracted by our approach with Term Frequency, Documents Frequency, TF-PDF and the weighting scheme described in [2]. This experiment validates the effectiveness of our method over these weighting schemes. In the second experiment, we demonstrate how we can identify and rank real-time hot terms effectively.

Experiment 1

Figs. 7 - 9 show the top 50 entities on Jan 04, 2007 in order from high to low by term frequency, document frequency and TF-PDF value respectively. Not surprisingly, most of the top terms are relatively common terms, such as *"U.S.", "Iraq", "Microsoft", "California"*, etc. These terms appear in news articles every day. There are two categories of these terms. One is the kind of term on which many news articles are produced, like "*Microsoft", "Google", "Saddam", "Bush"*. There are news articles on these terms almost every day.

Meanwhile, the other kind is more like containers, such as "*U.S.*", "*China*", "*California*", "*Los Angeles*". Events happening at these places tend to have these terms in the articles. Both kinds of these terms cannot be considered as hot terms.

Fig. 10 presents the top 50 ordered terms ordered by our weighting scheme. The result is much better than that derived by the other approaches shown in Figs. 7 – 9 and ref [2]. The relatively common terms are given low scores and none of them appear in our top 50 list. The genuinely hot terms extracted from hot articles rated by Yahoo! rank higher in our ordered list, such as "*ironport*", "*nardelli*", etc.

Experiment 2

Fig. 11 shows the "hotness" score change over time for the hottest terms "*iraq*", "*ironport*", and "*iphone*". "*iraq*" had a relatively low score over time since it is a common term and not considered as a "hot" topical term in our measurement. However, it gained a relatively higher score on Jan 09, 2007. The reason is that President Bush delivered an important speech on Iraq which drew a lot of attention and resulted in a lot of news articles about this topic.

Summary of results and comparison to other works

Compared *to tf, tf-isf*, and the term weighting scheme proposed in [2], our scheme can more accurately score and identify hot terms. Furthermore, our scheme models the short term burst and historical variation correctly. By comparing with real world data and other approaches, our system can extract hot news subjects effectively with higher recall. Furthermore, the result is relatively stable through continuous observation over one week.

VI. Conclusions

We propose herein a real-time hot news recommendation system to answer the question from users: "what is the hot news today?" In other words, given all news articles on the current day p and previous news archives, identify the hot news on day p.

We show two key characteristics of hotness: short-term burst and large historical variation. By incorporating the two important features of hotness, we also proposed a novel term weighting scheme for hotness to extract hot terms on given date. In order to compute the historical variation of a term correctly, we examined a lot of related factors and modeled the historical variation of a term by variation of interval (in day) between two occurrences of this term. With each news article on day p represented by hot terms it contains, the articles are grouped into clusters which represent hot events on that day.

We foresee the system to show on the internet in real time the hot topics over news feeds or other document corpora indicated, as shown graphically in the above figures. This will necessitate real time applications of the formulations we have developed to lead to the visualization shown here in, marking with different colors and icons the hot topics. The hot terms and hot new articles extracted are different from day to day, and reflect the hot events on each day.

References

[1] Seokkyung Chung, Dennis Mcleod, Dynamic topic mining from a news stream, In Proceedings of the 2nd International Conference on Ontologies, Databases, and Applications of Semantics (ODBASE 2003), Sicily, Italy.

[2] Gianna M. Del Corso, Antonio Gull′, Francesco Romani. Ranking a Stream of News, WWW 2005, May 10-14, 2005, Chiba, Japan.

[3] Kuan-Yu Chen, L. Luesukprasert, S.-c.T. Chou, Hot topic extraction based on timeline analysis and multidimensional sentence modeling, IEEE Transactions on Knowledge and data engineering, Volume 19, Issue 8, August, 2007.

[4] James Allan, Rahul Gupta, Vikas Khandelwal: Temporal Summaries of News Topics, SIGIR 2001.

[5] Jain, A.K., and Dubes, R.C.,1988, *Algorithms for Clustering Data*, Prentice Hall: New Jersey.

[6] Salton, Gerard and Buckley, C. (1988). "Term-weighting approaches in automatic text retrieval". Information Processing & Management

[7] Joel Larocca, Neto Alexandre, D. Santos, Celso A. A, Kaestner Alex, A. Freitas, Catolica Parana, Document Clustering and Text Summarization, 4th International Conference Practical Applications of Knowledge Discovery and Data Mining (PADD-2000)

[8] Khoo Khyou, Bun Mitsuru Ishizuka, Topic extraction from news archive using TF*PDF algorithm, Proc. Third Int'l Conf. Web Information Systems Eng, 2002.

[9] http://incubator.apache.org/uima/

[10] http://alias-i.com/lingpipe/

[11] Lan You, Xuanjing Huang, Lide Wu, Hao Yu, Jun Wang, Fumihito Nishino, Exploring Various Features to Optimize Hot Topic Retrieval on WEB, ISNN (1) 2004, 1025-1031

[12] Malu Castellanos, HotMiner: Discovering Hot Topics on the Web". Proceedings of SIAM workshop on Text Mining, Arlington, Virginia, April 2002.

[13] http://en.wikipedia.org/

Term	Definition
Term	Words and compound words (but not phrases) that are used in specific contexts. [13]
Entity	Atomic elements in text such as the names of persons, organizations, locations, expressions of times, quantities, monetary values, percentages, etc. Each entity has its name and is referred to by a term. [22]
News	News is any new information or information on current events which is presented by print, broadcast, Internet, or word of mouth to a third party or mass audience. [22]
Topic	A broader category of events, or a class of events, such as "volcanic eruption" [4]

Event	Some unique thing that happens at some point in time, such as “the eruption of Mount Pinatubo on June 15th, 1991” [4]

Fig. 1. Term definitions[A2]

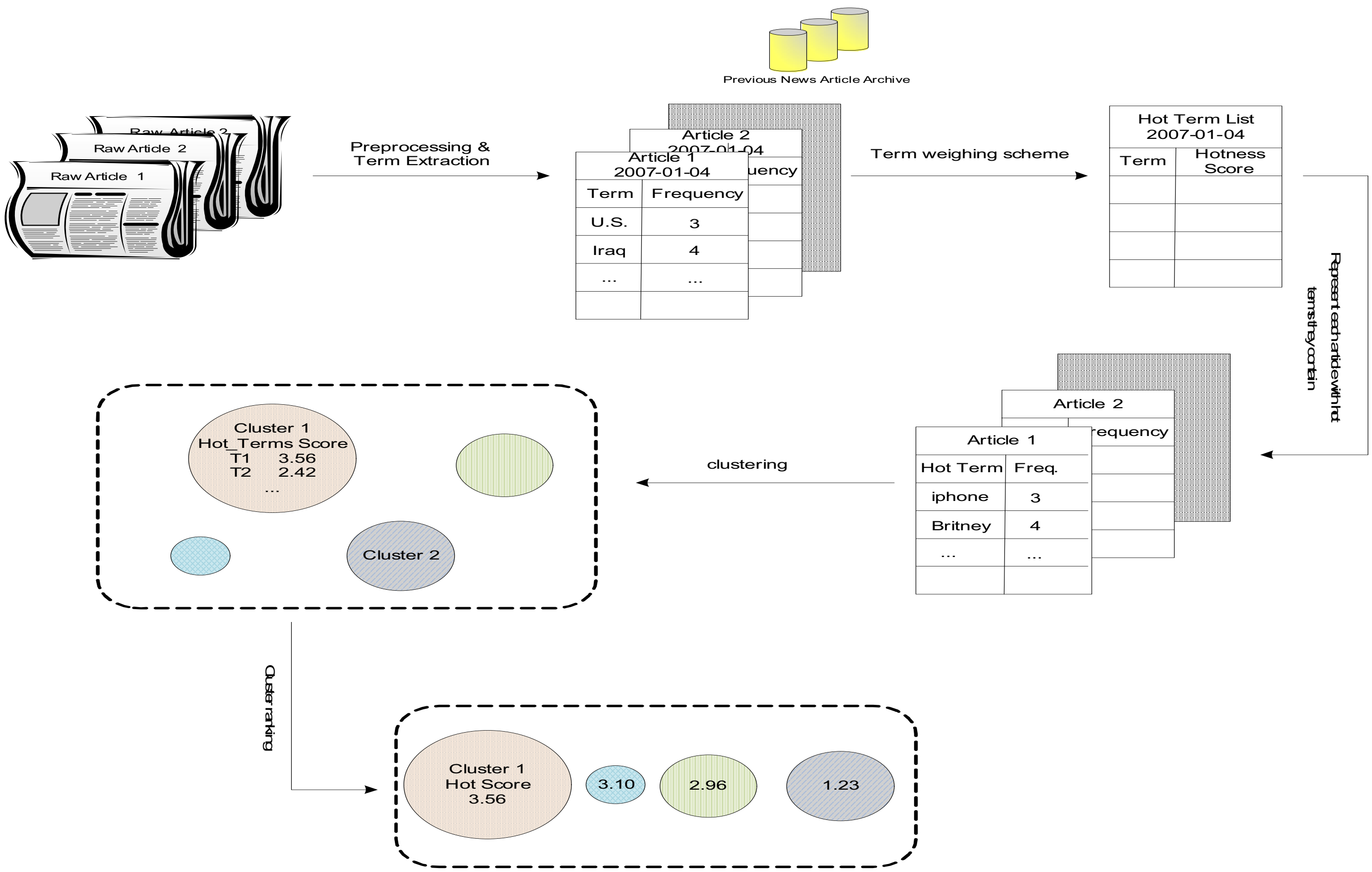


Fig. 2. Overview of our real-time hot news extraction system

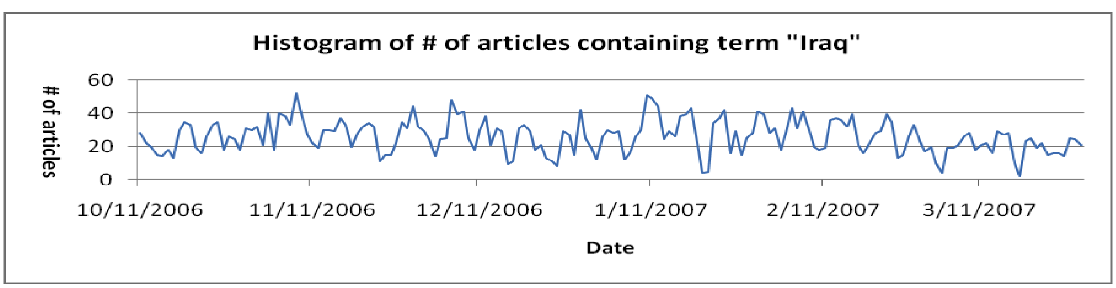


Fig. 3. Histogram of # of articles containing term “Iraq”

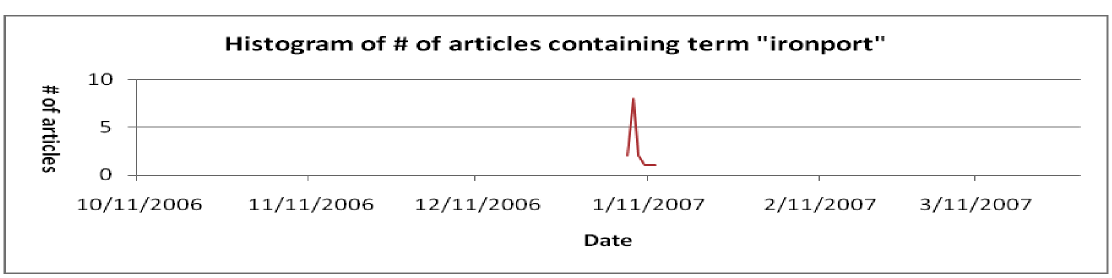


Fig. 4. Histogram of # of articles containing term "ironport"[A3]

Hot Term	Hotness Score	Frequency	Term Weight
term1	3.22	3 /(3+4+2+3)	3.22 * 3 / 12
term2	3.01	4/ (3+4+2+3)	3.01 * 4 / 12
term3	1.98	2/ (3+4+2+3)	1.98 * 2 / 12
term4	1.45	3/ (3+4+2+3)	1.45 * 3 / 12
…			

Fig. 5. Our representation of an article after hot terms are extracted

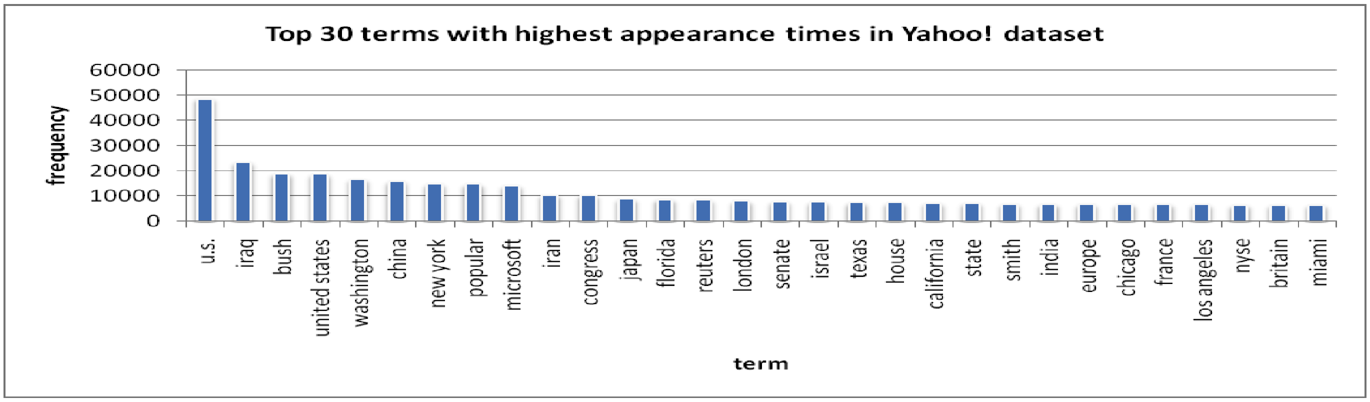


Fig. 6. Top 50 terms with highest appearance times in Yahoo! dataset

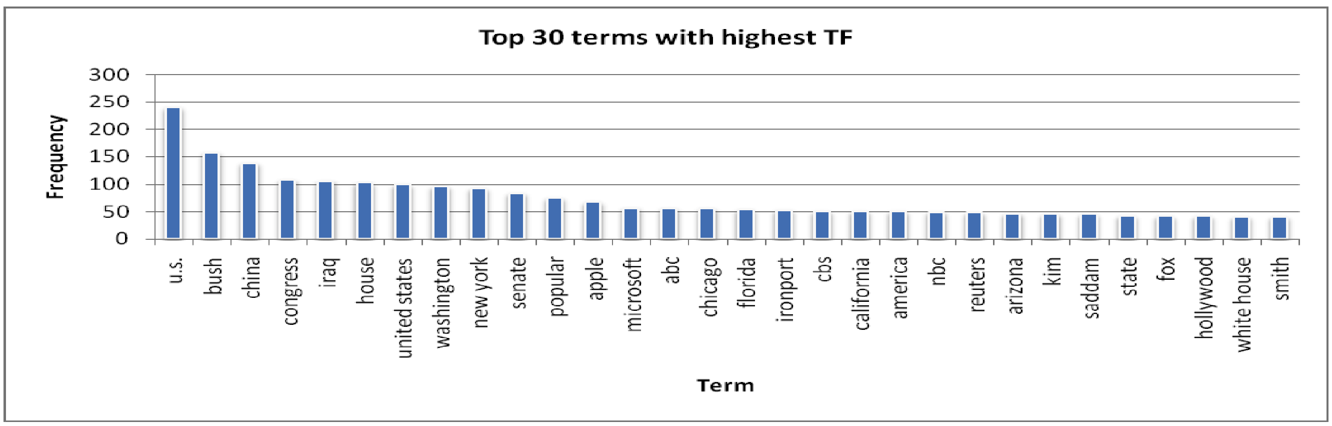


Fig. 7. Top 50 terms with highest TF[A4]

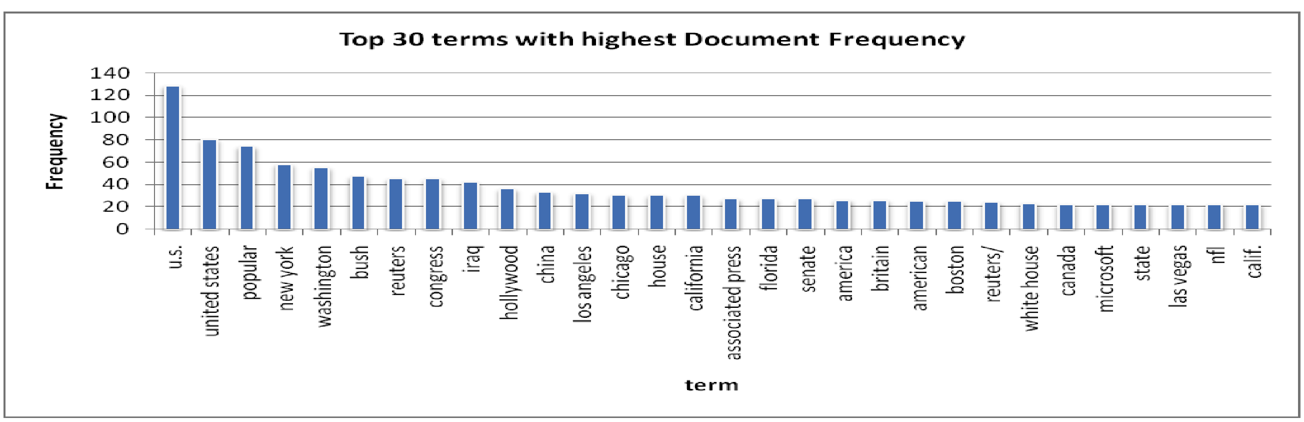


Fig. 8. Top 50 terms with highest Document Frequency[A5]

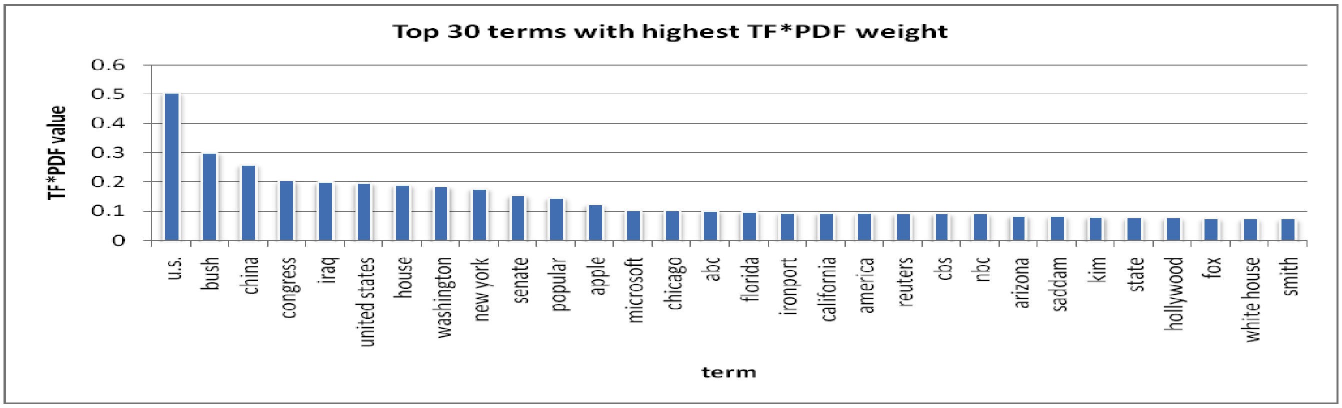


Fig. 9. Top 50 terms with highest TF*PDF weight

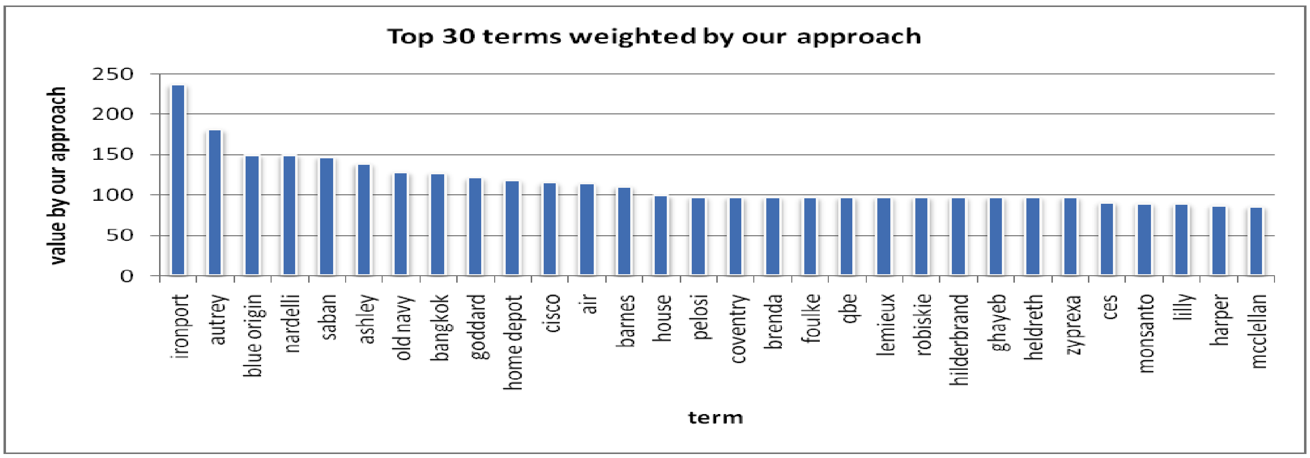


Fig. 10. Top 50 terms weighted by our approach

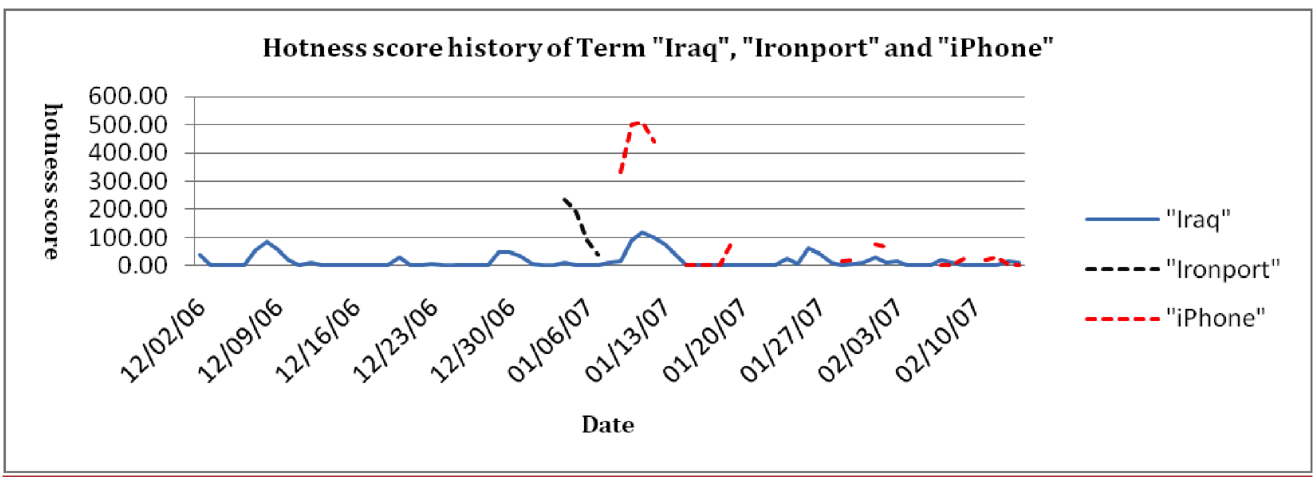


Fig. 11. Hotness score history of three hottest terms

Visual Analysis of Spatial Data through Maps of Chorems

Davide De Chiara[1], Vincenzo Del Fatto[1], Robert Laurini[2],
Monica Sebillo[1], Giuliana Vitiello[1]
[1]*DMI, Università di Salerno, Italy*
{ddechiara, delfatto, msebillo, vitiello}@unisa.it
[2]*LIRIS, INSA de Lyon, F-69621 Villeurbanne*
Robert.Laurini@insa-lyon.fr

Abstract

Chorems were initially introduced by the geographer Brunet in 1986 to visually synthesize a territory and a set of phenomena concerning it. A recent collaboration among research institutions has adopted the chorem concept aiming at providing a new cartographic solution capable to support analysts' work. Current results have illustrated how chorematic maps can be used to summarize the content of geographic databases and a system prototype has been specified to produce them in a semi-automatic way. In this paper we aim to demonstrate that chorems can be exploited to both catch a thematic global view of a territory and its phenomena, and investigate a single phenomenon by accessing data characterizing it. The proposal is based on Shneiderman's mantra for designing human interfaces "Overview, zoom and filter, details on demand", and an example is described to show how functionalities work.

1. Introduction

One of the major challenges the GI community is facing is to provide decision makers with advanced tools capable of semantically and visually integrating quantitative, qualitative and cognitive aspects of a domain of interest. In fact, when a large amount of data is available, synthesizing information and deriving insight from it may result a time-consuming and expensive task. Analogously, to discover hidden information and perform specific analytical tasks may require application domain expert users and high performance tools.

A significant contribution to this issue may be provided by the discipline of Geovisual Analytics (short for Geospatial Visual Analytics), whose aim is to maximize human capacity to perceive, understand and reason about complex and dynamic geodata, by supporting the analytical reasoning process [9]. In fact, the adoption of visual interactive methods from the geovisualization research, and their integration with new possibilities offered by computational techniques, enhance the real contribution that this discipline could offer to data exploration and decision-making processes, due to its capability of combining geospatial information with "human vision and domain expertise" [4, 6, 7].

The research we are carrying out along this line aims to provide a new cartographic solution capable to support analysts' work in representing dynamics, movements and changes that underlie possible problems. The proposed solution is based on the *chorem* and chorematic map concepts. The former corresponds to a schematic territory representation which eliminates details not useful to the map comprehension [1]. The latter is an immediate synthesis of data of interest and summarizes scenarios involving static objects and dynamic phenomena by associating them with schematic visual notations. In Figure 1 a set of chorems are depicted, namely industrial zones and big poles, which make up a chorematic map referring to the French territory.

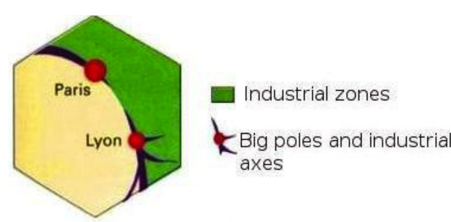


Figure 1. An example of a chorematic map

The potentialities of chorems have been initially presented by Laurini *et al*. in [5], where the authors list different roles that chorems may play in supporting expert users in daily activities. The authors state that chorems can be used to represent geographic knowledge, to visually summarize database contents and finally to underlie the creation of a novel entry system for geographic databases.

The aim of this paper is to enhance the role that a chorematic map may play in geographic domains, by extending the semantics associated with it through a more expressive visual notation. In particular, we aim to demonstrate that a set of chorematic maps deriving from a unique spatial database may be used to both catch a thematic global view of a territory and its phenomena, and investigate a single phenomenon by accessing data characterizing it.

In order to reach this goal, we propose to adopt Shneiderman's *mantra,* namely "Overview, zoom and filter, details on demand" [8], and assign these operations with a specific behaviour.

Namely, our approach is meant to allow users to acquire information from the underlying database by interacting with chorematic maps which visually summarize its content. Each task of interaction assumes a context-sensitive meaning and invokes a proper function among the ones specified in agreement with the *mantra*. As an example, when a zoom / filter combination is applied on a chorem, users are provided with data from spatial dataset which initially contributed to its definition. In particular, if the generation of the selected chorem has been determined as an aggregation of simpler concepts, then a proper zoom may invoke their visualization both in terms of chorems and aggregated data.

The paper is organized as follows. Section 2 recalls the chorem classification in terms of territory and associated phenomena. Section 3 introduces the proposed method for accessing and analyzing spatial data. In Section 4 an example shows the specified functionality applied to a chorematic map. Conclusions are drawn in Section 5.

2. Chorem classification and ChorML structure

As for the chorem capability to visually summarize database content, in [3] the results concerning this aspect are detailed. In particular, the authors provide a definition and a classification of chorems meant both to homogenize chorem construction and usage and to provide a usable framework for computer systems.

Chorems are formally defined as follows:

– Proto-chorems: Data items, properly cleaned and organized, on which SQL queries and (spatial) data mining functions can be applied, in order to discover patterns.
– Chorem element: Each basic graphic element that may represent either a single geographic object or a phenomenon.
– Chorem: Set of chorem elements of the same typology.
– Chorematic Map, Set of chorems which schematizes data of interest related to a place,

Moreover, chorems are classified into three main categories:

– Geographic chorems represent objects with simple spatial features, such as points, lines, polygons.
– Phenomenal chorems describe spatio-temporal phenomena involving one or several geographic chorems. The initial set of phenomenal chorems we have identified consists of three types, Flow, Tropism, and Spatial Diffusion.
– Annotation chorems represent map labels.

In [3] a construction process of chorematic maps is also specified by showing the tasks both for extracting relevant information from a large amount of data about territory and geographic phenomena, and for representing them in a form suitable for consumption by the ultimate users. The whole process has been specified in terms of components and algorithms. As depicted in Figure 2, this process is accomplished by a system which performs two main tasks, namely chorem extraction and chorem visualization. Moreover, an XML-like language, named ChorML, has been specified to store information about chorems. It consists of three levels which differ in terms of structure and typology of stored information.

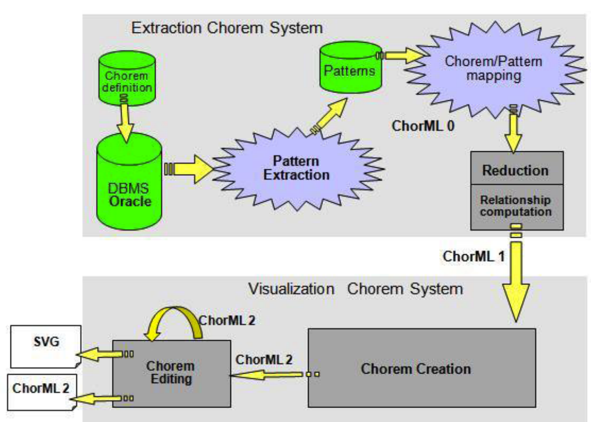


Figure 2. The Chorem extraction and visualization System

In particular, level 0 of ChorML is made up of XML and GML tags. As shown in Figure 2, it supports the chorem extraction by storing information about proto-chorems, origin of data (lineage) and functions applied to them in order to obtain a list of chorems. Figure 3(a) illustrates the level 0 structure.

Level 1 of ChorML is still a combination of XML and GML tags. It connects the chorem extraction and chorem visualization by specifying the results of the application of (spatial) data mining algorithms. In particular, as depicted in Figure 3(b), the items of the language are:

- general information containing id, project name, author name, creation date, layout, reference system, original database name, last update;
- the chorem list where geographic data are GML coded;
- pre-legend containing a description of chorems;
- spatial relationships between chorems.

Currently, this level is totally specified with an XML grammar presented in [2].

Level 2 of ChorML corresponds to the code imported /exported by the chorem visualization sub-system in terms of XML and SVG tags. The elements of level 1 and level 2 differ from the coding of geographic components, GML and SVG respectively. Figure 3(c) illustrates the level 2 structure.

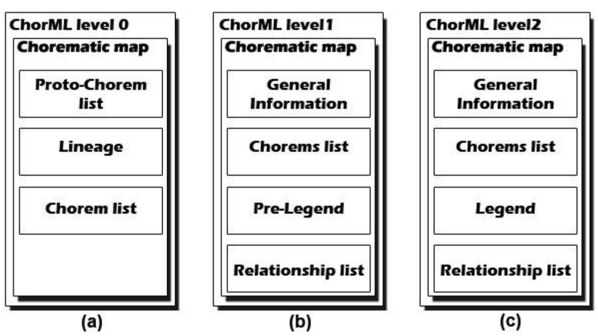


Figure 3. ChorML levels.

3. Accessing Spatial Data through Maps of Chorems

The idea underlying our proposal is based on the "Overview, zoom and filter, details on demand" approach, i.e. macroscopic versus microscopic. In agreement with this approach, we assign each interaction task with a specific meaning, aiming at satisfying each type of user's requirement.

In the following subsection we describe the operations we propose, meant to allow users to interact with a chorematic map in order to perform visual analytics tasks, and the extension of ChorML structure capable to support the introduced functionality.

3.1 Zooming and filtering a chorematic map

The first step envisaged by the Shneiderman's mantra is the *overview*. It is meant to obtain a global view of a situation featuring a domain of interest. In our proposal, looking up a map may exhaustively correspond to an overview. By it, users can both obtain an immediate synthesis of the mini-world under investigation and capture salient information.

As for the *zoom and filter*, the goal is to focus attention either on a reduced portion of the whole space or on details characterizing a feature.

In our proposal we pursue a similar target. By zooming and filtering, users can gradually reduce the search space and select a subset of data. However, given the implicit meaning of the concept of chorem, which is meant to convey information about even complex spatial patterns, zooming will mean using different geographic scales or thematic disaggregation, whereas filtering will reflect conditions and criteria. In particular, four basic operations are defined in our proposal, so that users are able to navigate the chorematic map from an initial overview to a particular detail, by using a combination of them. In this paper, such operations are referred to the geographic chorems and to a particular phenomenal chorem, namely the Flow. These operations are described as follows:

- Geographic zoom. This operation corresponds to the traditional GIS zoom operator. By it, the user can visualize the map at different scale levels.
- Semantic zoom. This operation allows the user to analyze chorems in detail, without affecting the map scale. It differently breaks a chorem into sub-components, depending on the group it belongs to. In case a geographic chorem is involved, this latter is simply split up. When a semantic zoom is applied on a flow, both the flow and the geographic chorems related to it are decomposed.
- Geographic filter. This operation allows user to visualize only the phenomena related with the visible geographic chorems, i.e., it corresponds to a filter which takes in account the phenomena concerning the visualized portion of the map.
- Semantic filter. This operation allows the user to filter chorem elements which satisfy a particular condition. The user may use a threshold applied to the descriptive component in order to reduce the visualized data

The third step of Shneiderman's *mantra*, namely *details on demand*, is meant to obtain detailed information about a particular map object or phenomenon. Our approach allows the user to select and query a chorem element in order to obtain descriptive information related to it, stored in ChorML.

3.2 The extended version of ChorML

In order to demonstrate that a chorematic map may be used to investigate phenomena by accessing data through the previously described operations, an extended version of ChorML is introduced. In particular, it is necessary to operate during the construction phase of a map by adding ChorML *ad hoc* information, useful to perform visual analytics tasks. In fact, the ChorML currently stores information about the visual summary. No information about the initial data source is available. Then, a revised version of ChorML is necessary, which collects both data referring to detecting phenomena, in terms of spatial and semantic component, and functions used for aggregating those data. A multiscale structure is proposed to store within the ChorML intermediate data processed during the map construction. In particular, spatial components are stored as geometric primitives whereas the associated alphanumeric data can be gained by specifying dynamic views used for their aggregation.

To this aim, we propose new hierarchical tag instances, where information about multi-scale representations can be stored. In particular, Figure 4 shows the original structure of ChorML, in terms of chorem list and elements, whereas Figure 5 illustrates ChorML tags associated with each chorem element.

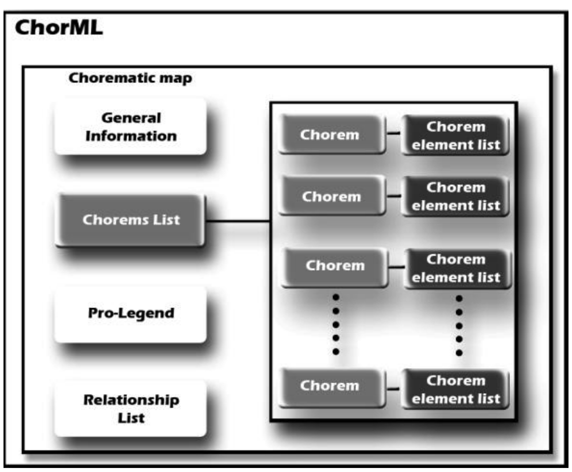


Figure 4. ***The original structure of ChorML***

The structure we propose is depicted in Figure 6. It extends each chorem element by a tag representing a list, thus creating a hierarchy of nested chorem elements. This organization corresponds to a multi scale representation, which allows to access the different levels of aggregation of the original data. In fact, once stored, such information may be managed to answer queries that end users pose by spatially manipulating a chorematic map. Selecting a chorem, zooming a portion of it, filtering region of interest, represent actions which invoke the previously defined operations.

In order to show how users may acquire information about a phenomenon of interest, in the following some examples of interaction tasks are illustrated.

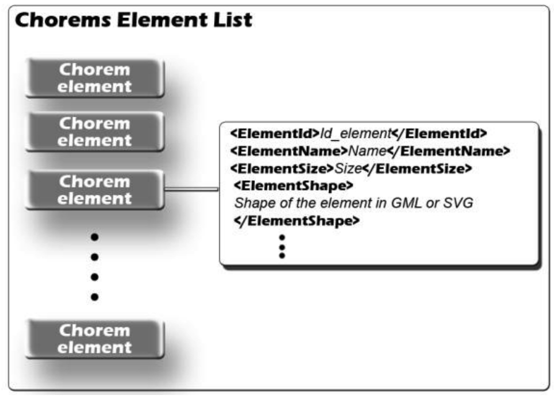


Figure 5. ***ChorML tags of a chorem element***

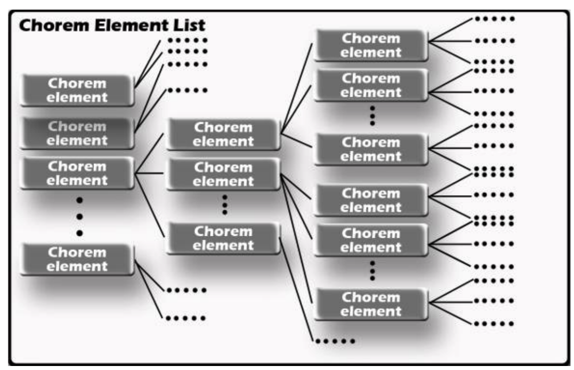


Figure 6. The extended structure of ChorML

4. Analyzing the flow of migrating population among macro-regions in Italy

They are meant to describe the usage of the previously described operations in a specific domain along with the focus on the corresponding level of ChorML. Figure 7 shows a chorematic map built by the system proposed in [3]. It represents five macro-regions and the three most significant flows of the Italian migrating population among them.

The former set belongs to the geographic chorem group, the latter are instances of the phenomenal chorem group.

Figure 7. A chorem map under investigation

The construction of the map started from data reported in Table 1, namely a set of demographic data related to the Italian regions. By clustering and summarizing them according to proper functions, data reported in Table 2 have been obtained, which synthesize the number of displacements among the five recognized macro-regions. Finally, by applying a threshold of 25000, the three most important flows have been highlighted.

Table 1. Demographic data about regions

Shape	Region	Piemonte	V_d_a	Lombardia	Taa	Veneto
Polygon	ABRUZZO	235	4	792	77	337
Polygon	BASILICATA	377	6	769	42	166
Polygon	CALABRIA	2244	167	5308	246	1083
Polygon	CAMPANIA	2648	69	8591	508	3183
Polygon	EMILIA-ROMAGNA	852	55	4168	266	1953
Polygon	FRIULI-VENEZIA GIU	234	13	747	130	1774
Polygon	LAZIO	1230	55	3701	376	1968
Polygon	LIGURIA	3184	77	2274	109	379
Polygon	LOMBARDIA	5595	128	0	726	3579

It is worth to note that differently from the previous approach, which requires to store in summarized data reported in Table 2, the current approach needs to handle starting data of Table 1. This allows for managing data at different levels of aggregation, even not planned during the construction task.

In the following, we exemplify the operations previously defined, which allow users to browse the map from a global sight to a detailed feature, in agreement with Shneiderman's *mantra*.

Table 2. Displacements among macro-regions

	North	Center	South	Sicily	Sardinia
North	0	18112	31130	13150	5248
Center	20903	0	10594	2084	1428
South	66668	22132	0	2521	847
Sicily	26489	4172	2483	0	319
Sardinia	6323	1420	588	269	0

Let us consider the following request, namely to determine the region which received the highest number of migrating people from the South of Italy, disregarding the islands.

By recalling the *overview* step, five macro-regions and three arrows, representing the most relevant migration flows, can be detected as shown in Figure 7. In particular, the arrow thickness is visually proportional to the number of people who migrated between the two macro-regions.

The *zoom and filter* step is then useful to determine the most significant flow from the South macro-region towards to a specific Northern region. In fact, by invoking a geographic filter, only the two flows depicted in Figure 8 are selected, the third one is rejected because it involves the Sicily macro-region which is not shown in the map.

Figure 8. The application of a semantic filter

Then, in order to obtain the most relevant migration flow, as shown in Figure 9, a semantic filter on the phenomenal chorem, depicted in Figure 8, is applied.

Figure 9. The application of a semantic filter

Moreover, by applying a semantic zoom on the destination of the detected flow, the user decomposes it, displaying regions featuring the macro-region chorem, accordingly.

This operation allows user to understand how the previous flow was created by aggregating components of a lower level of representation, as depicted in Figure

11. In case a more detailed view is required, a geographic zoom can be used in order to focus the user's attention on a portion of the whole map.

Figure 11 shows the application of such an operation on the Northern regions.

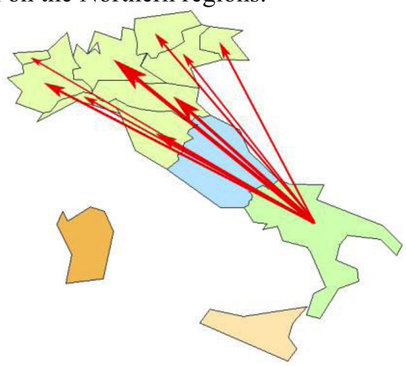


Figure 10. The application of a semantic zoom

Finally, by applying again a semantic filter, the most significant migration flow is selected.

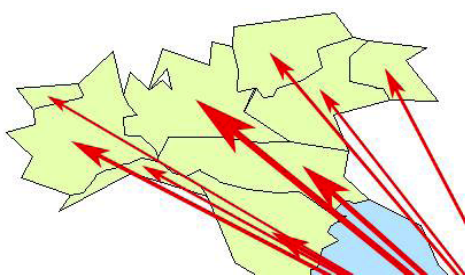


Figure 11. The application of a geographic zoom

In particular, the third step, *details on demand*, may be performed in order to obtain the exact number of people associated with the flow, namely 568 people, as shown in Figure 12.

Figure 12. The final result of the running example

5. Conclusions

In this paper we have illustrated how the visual notation associated with the meaning of a chorematic map may be handled to enhance the role it plays in a geographic domain. Some basic operations have been specified which manipulate geographic and semantic components of chorems in order to derive information underlying the map.

In the future we plan to extend our proposal to embed also other types of chorem groups, such as tropism and diffusion. Moreover, a usability study is also scheduled to evaluate the effectiveness of our proposal.

6. References

[1] Brunet R. La carte-modèle et les chorèmes. *Mappemonde* 1986; 4: 4–6.

[2] Coimbra A. (2008) "ChorML: XML Extension for Modeling Visual Summaries of Geographic Databases Based on Chorems", Master Dissertation, INSA-Lyon, June 2008.

[3] Del Fatto V., Laurini R., Lopez K., Sebillo M., Vitiello G., "A chorem-based approach for visually synthesizing complex phenomena", in *Journal of Information Visualization*, vol. 7, n. 3, 2008, pp. 253-264. Andrienko G., Andrienko N., Dykes J., Fabrikant S. and Wachowicz M. (Eds). Palgrave Macmillan Journals.

[4] Jiang, B., and Li, Z. 2005. Editorial: Geovisualization: Design, Enhanced Visual Tools and Applications. *The Cartographic Journal*, 42(1), pp. 3-4

[5] Laurini R., Milleret-Raffort F., and Lopez K., A Primer of Geographic Databases Based on Chorems. Workshop on Reliability in Decentralized Distributed Systems (RDDS) LNCS 4278, pp. 1693 – 1702, 2006. © Springer-Verlag Berlin Heidelberg 2006.

[6] MacEachren, A.M. and Kraak, M.J. 1997. Exploratory cartographic visualization: advancing the agenda. *Computers & Geosciences*, 23(4), pp. 335-343

[7] MacEachren, A.M. 2004. Geovisualization for knowledge construction and decision support. *IEEE Computer Graphics and Applications*, 24(1), pp.13-17

[8] Shneiderman B. (1997b) "*Designing the User Interface*." Third edition. Addison-Wesley. Publishing Company. 600pp.

[9] Illuminating the Path – The Research and Development Agenda for Visual Analytics, Edited by J.J. Thomas and K.A. Cook. National Visualization and Analytics Center TM. IEEE 2005.

SOFTWARE VISUALIZATION USING A TREEMAP-HYPERCUBE METAPHOR

Amaia Aguirregoitia [1] , J. Javier Dolado [2] and Concepción Presedo [1]
[1] University of the Basque Country,La Casilla 3, 48012 Bilbao (Spain).
amaia.aguirregoitia@ehu.es, conchi.presedo@ehu.es
[2] University of the Basque Country,Pº Manuel de Lardizabal, 1, 20018 Donostia(Spain).
javier.dolado@ehu.es

Abstract

This paper presents T-Cube, a new graphical representation model for controlling and managing the processes of software project development. T-Cube uses metaphors and visual representation techniques to explore several key indicators in order to support problem detection and resolution. The resulting visualization addresses diverse management tasks, such as tracking of deviations from the plan, analysis of patterns of failure detection and correction, overall assessment of change management policies, and estimation of product quality. *This structured approach uses a metaphor with the Rubik-Cube to deal quickly and accurately with visual queries. It presents a hypercube composed of different faces. Each face contains a TreeMap to represent the software development process information. Finally, the paper shows the implementation of the tool in Java with data of a real project and the results of testing the tool with users attempting several information retrieval tasks. User response time and efficiency were recorded and the conclusion analyzes efficiency and response times using the T-Cube visualization system. The utility of the tool was positively evaluated.*

Key Words
Visual Information Systems, Visualization Software, Software Visualization, Software Project Management.

1. Introduction

Management of software development processes is a difficult task. Often projects are behind schedule and the resulting applications lack quality. Managers require different types of data, for instance written reports from project managers or software metrics like productivity, quality, adherence to schedule and budget. To assist in exploration and analysis of the high volumes of data required, our research focuses on the design of a tool to efficiently process visual queries on the key measures for software development management.

The next section presents the related work. Then a description of the tool is given. In the conclusion, we assess our work and present the major benefits of the proposed visualization.

2. Related work

2.1. Definitions

Before presenting our findings, it is necessary to provide definitions for some relevant concepts. The term entity refers to any distinguishable object in the empirical world for which a measurement can be applied [1]. In our research, the tasks of the software development process are the entities to be measured. The term attribute refers to the property of an entity that can be determined quantitatively, that is, for which a magnitude can be assigned. The measurement of an attribute of an entity is the characterization of that attribute in terms of numbers or symbols. We adopt this general definition, which is consistent with most of the standards and with the literature.

2.2. Software visualization

Prior research on software visualization has mainly focused on the representation of the technical implementation aspects of the project. More specifically, it has represented source code or data structures with the primary objective of understanding and improving the code. Animation of program behaviour, presentation of relationships among functions and interactions for better comprehension or computation visualization [2] are some other threads of activity. Source code changes or architecture aspects are frequently represented using running time diagrams indicating function or module calls. Some other representations include graphs, statistics or relatively traditional diagrams representing certain code or performance aspects [3] and most of them have the primary objective of improving the code [4].

Augur [5], a comprehensive visualization system for CM repositories of source code, contains multiple, dynamic views of data in an attempt to unify views of “activities” and “artifacts” and offers project’s source code, author network, and other development statistics. The system StarGate [6] is an information visualization project that uses different techniques to assist the user in understanding the complex interactions between developers and software. It visualizes the code repository and social network of developers associated with a software project in one integrated representation.

2.3. Software project management visualization

The idea of measuring processes and their outcomes plays an important role in software engineering activities. Much research has been dedicated to designing software metrics, initiatives for software process development improvement and decision making [7] and there are a number of experiences reporting the problems in the measurement area. To mention but a few , most software engineers and managers did not use the measurement program, the data suppliers did not understand the definitions of metrics and measures, the project indicators focused on resources and deadlines and ignored the quality of software and few people were capable of interpreting the presented data [8].

Besides the measuring difficulties, an additional complexity of the project management and control area is how to present and interpret the information. At present, Gantt charts are predominantly used for the mapping of projects in organizations. While they are effective for planning a project they are not effective for communication purposes, especially when different groups are involved [9].

The UML package diagram is another frequently used graphical resource, as are some other formal and informal diagrams for the static view [10] (such as class or component diagrams). In addition to a static view, a functional view can be provided through a use-case diagram with color-coding, which offers a quick overview of the functionalities of the present development state. The functional view of status visualization is not an alternative to the static view, but a complementary view.

Some previous work in the field has led to the development of visualizations for software management but most of them address specific problems and are mainly focused on metrics of particular areas. One of them is Tarantula [11], which is useful for finding likely faulty sections of the code and gives the developer information about the results of a faulty program's execution on an entire test suite. Another example, SVAW [12] offers visual representations for assistance to human schedulers.

However, there are some other visualizations which are unknown and rarely used that have been reported as exceptionally powerful in the communication area as the tube map proposal in [13].

3. The treemap hypercube metaphor

3.1. The measures selected

The information presented using the metaphor came out as a result of the analysis of data from several works and surveys in the software project management area. What appears below summarizes the most relevant aspects found in these surveys that guided the definition of the measures in the system.

a) The importance of scheduling and project estimation tracking

According to these studies 26% percent to 34% of IT projects are failures (cancelled projects plus delivered projects with unsuccessful performance). In 2007, around 37 % of the people surveyed rated the ability to meet budget targets as fair or poor, and 47.10 % rated "meeting schedule targets" as fair or poor. The most critical performance problems in delivered software projects are therefore estimating the schedule and meeting that estimate [14].

b) The importance of requirements changes management.

According to a 2007 survey changes in requirements and scope were primary reasons for project cancellation chosen by 33 % of the respondents[14]. Some other studies stated that over half of the higher performing projects exhibited a higher capability in Requirements Development and Management, suggesting the value of effective requirements management in producing successful project outcomes.

c) The relevance of risk identification and analysis.

Software projects have long been recognized as high-risk ventures prone to failure [15] and risk management has become recognized as a best practice in the software industry for reducing the surprise factor [16]. By observation of different projects it has been proven that the most successful project managers are good risk managers [17]. A formal risk management process provides multiple benefits to both the project team and the development organization as a whole. Risk management in software projects is important to avoid disasters, avoid rework, focus and balance effort and stimulate win–win situations [18].

d) The benefits of failure identification, classification and correction.

A key aspect of producing reliable software is the analysis of failure distribution by phase (requirements, analysis, design and documentation) as proposed in the failure types defined in [19]. This analysis implies the existence of data acquisition procedures and technical reviews - evaluations of a software product by a team of qualified personnel to determine its suitability for its intended use and to identify deviations from specifications and standards- that could, as an example, be classified into requirements, analysis, design and documentation reviews according to the proposal in [20]. A list of measures for each task involved in the development process is presented in *Table1*.

Table 1.- List of measures
Task effort
Estimated task effort
Task cost
Planned task cost
Number of Requirement failures
Number of Design failures
Number of Code failures
Number of Documentation failures
Number of other type failures
Number of total failures
Number of failures detected by the client
Number of failures detected by the developers
Requirement failure detection effort
Design failure detection effort
Code failure detection effort
Documentation failure detection effort
Other type failure detection effort
Failure detection total effort

Number of Requirement reviews
Number of Design reviews
Number of Code reviews
Number of Documentation reviews
Number of other type reviews
Number of total reviews
Requirement failure correction effort
Design failure correction effort
Code failure correction effort
Documentation failure correction effort
Other type failure correction effort
Failure correction total effort
Number of changes required
Number of changes rejected
Number of changes implemented
Number of changes pending
Number of deliverables planned
Number of deliverables rejected by the client
Number of deliverables accepted by the client
Number of pending deliverables
Number of detected risks (with description and type)
Effort deviation
Cost deviation
Risk detection effort

Once the information for each task has been defined, one of the questions that arise is how to arrange tasks.

3.2. Treemaps and Metaphors

In our proposal we use both metaphors and treemaps. A Treemap is a space-constrained visualization of hierarchical structures. It is very effective in showing attributes of leaf nodes using size and colour coding. Treemap enables users to compare nodes and sub-trees even at varying depth in the tree, and help them spot patterns and exceptions. The Treemap was first designed by Ben Shneiderman during the 1990s. It is extremely efficient to represent extensive attributes (sizes, costs, value) of elements arranged in a hierarchy In a treemap each node has a name and an associated size. The size of the leaves may represent, for instance, the effort of individual tasks, the size of non-leave nodes is the sum of the sizes of its children. The treemap is constructed via recursive subdivision of the initial rectangle. The size of each sub-rectangle is proportional to the size of the node. The direction of the subdivision alternates per level: first horizontally, next vertically, and so on. As a result, the initial rectangle is partitioned into smaller rectangles, such that the size of each rectangle reflects the size of the leaf. The structure of the tree is also reflected in the treemap, as a result of its construction. Colour and annotation can be used to give extra information about the leaves [21]. This technique has been applied to a wide variety of domains: to present large number of images grouped by directory, to analyze file systems, financial analysis [22] or sports reporting [23].

Metaphors are an important tool in information visualizations as they provide familiar cognitive models to help users to browse unfamiliar information spaces [24]. Six advantages of Visual Metaphors have been described in previous works: (1) to motivate people, (2) to present new perspectives, (3) to increase remembrance, (4) to support the process of learning, (5) to focus attention and support concentration of the viewer, (6) to structure and coordinate communication [25].

Most of us played with a Rubik Cube when we were children, so we use this metaphor as a way to leverage its familiarity to enable users to better understand the tool and how it organizes tasks. In a Rubik cube, the six faces are covered by 9 stickers of six solid colours. A pivot mechanism enables each face to turn independently, thus scrambling the colors. For the puzzle to be solved, each face must be a solid color.

The purpose of using the metaphor with the Rubik cube is to assist in the comprehension of the task structure and rotation mechanism of the visualization even when there are apparent discrepancies. The Rubik cube is composed of smaller cubes and in order to solve it one needs to do two things: rotate the cube to see the different sides one by one and analyze colors. In the visualization, the project is divided into tasks (comparable to the small pieces or squares of the cube) which are organized into different facets (comparable to the faces of the cube) and only one of the sides is visible at a time while the rest are hidden. Each task has attributes that define the size and color of the corresponding square (as color in Rubik cube pieces) and the tool utilizes rotation to allow access to the other sides. Two more variables affect the T-Cube visualization: the criteria selected to arrange tasks on one or another side of the cube (different categories of tasks) and an attribute to be used to determine the colour and size of the task.

3.3. The Treemap Hypercube metaphor for Software management

There are a lot of tasks and measures about project tasks that we would like to be able to arrange, group, and move as with a Rubik Cube. Furthermore, a Rubik Cube inherently hides information when you focus on one face of the cube, which is an interesting feature when analysing only a cluster of tasks. The overall layout of how the metaphor is applied is presented in *Figure 1*.

The first step of the analysis is to define what characteristic of the tasks will define the clusters or facets of the hypercube. The criterion selected defines the number of sides of the hypercube and therefore, the list of available faces in the upper right part of the screen. In the example, when “Project phase” is selected the system presents one facet of the hypercube at a time using the previously defined phases of the project and it puts together in a face the tasks corresponding to “Design” phase, in another side the ones related to “Testing” phase and so on.

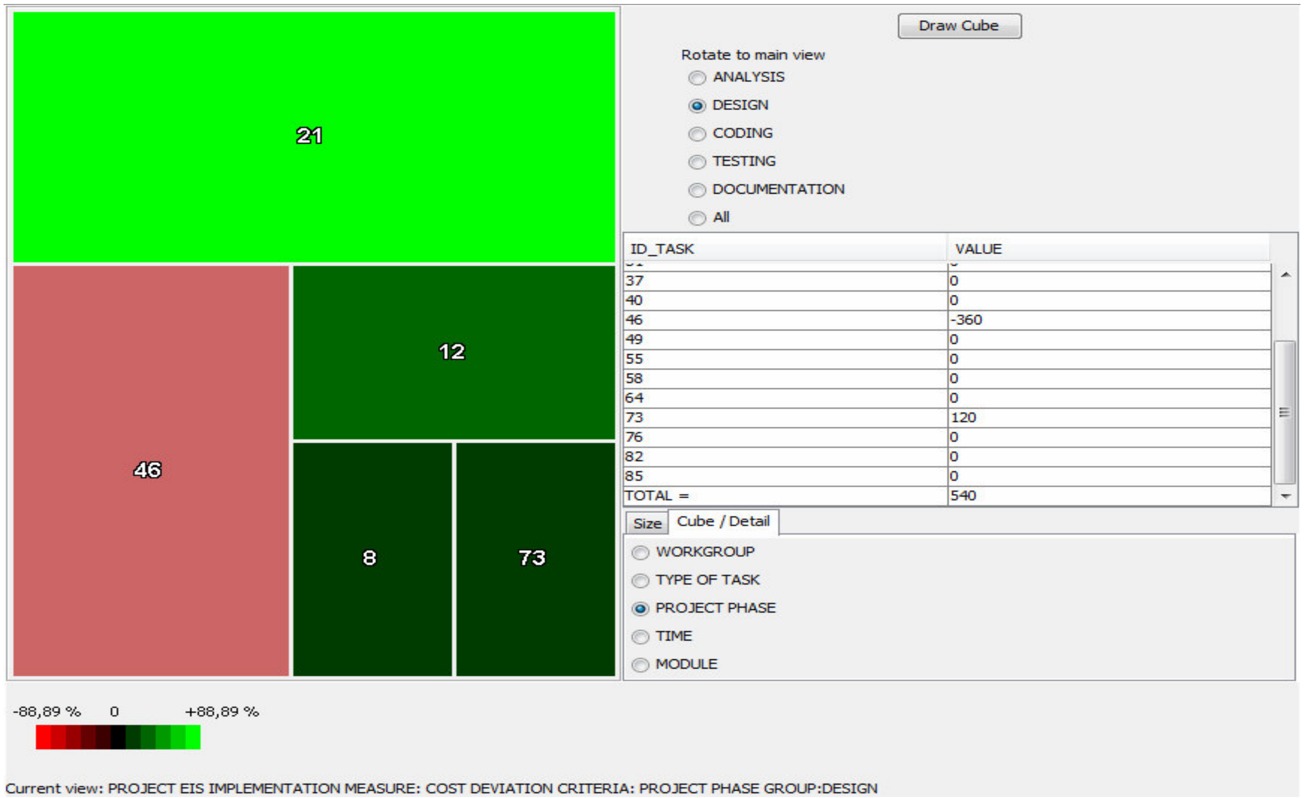


Fig. 1. Overview of the screen

The tasks in the example can also be arranged on faces using any of the following characteristics or criteria: workgroup, type of task, module or time. As an example, if the selection is workgroup, each face would contain all the tasks performed by the same workgroup and the hypercube would have as many faces as workgroups. The lower right part of *Figure 1* presents the radio buttons with the different criteria for defining the sides of the hypercube, while the upper part shows the side selection and rotation method.

The information in the main view is presented using the treemap space-filling methodology. In a typical treemap a square can represent either a leaf node or a group of items and the user moves from a group level to its leaves and vice versa by clicking on the square. In T-Cube, squares represent only the leaf nodes of the facet currently in the main view and the user is obligated to make radio button selections to modify the grouping criteria or to change from one group to another. However, in a typical treemap, clicking on the squares makes all these operations possible. This design, which is consistent with the metaphor, has the objective of assisting the user in structuring the information into criteria and groups and keeping this structure in mind.

Once the cube has been defined, the user chooses what face is to be displayed. The user selects from the list one of the facets, clicks on "Redraw cube," and the cube rotates, giving access to the side the user wants to analyze, what we call "Rotate to main view".

For an overall view, the system lets the user select the "Show All" option, which presents the information of all the tasks in a single view. It is worth noting that also a time based cluster creation and visualization is available. The flexibility for task arrangement of the tool is extremely useful since it enables to assess software from very different perspectives. The cube changes when a different criterion is selected and the visualization changes when the cube rotates as a facet is selected.

It is worth noting that any of the indicators described in *Table 1* can be selected to define the size of the presented tasks in the treemap. That means that this simple visualization offers a very complete set of data and that it is valuable for studying problems of different nature and complexity in the development process. Planning deviations tracking, analysis of defect detection and correction patterns, overall assessment of change management policies, evaluation of deliverable status, or product quality estimation are some of the problems that can be addressed by selecting different measures. In the lower part of the screen there is a text to inform the user about the currently visible data, which shows the project name, the measure selected from the list, and the criteria and group currently in the view.

The values of the selected measure for each task in the main view are presented using a table accessible from the Cube/Detail as shown in *Figure 1*. This table complements the visual information and it presents also the sum of all the tasks for the selected variable. This information is really valuable since the Tremap offers only individual data as a percentage of the group. The use of colour and size lets the user identify and focus on troublesome areas easily and the treemap allows the user to apply visual filters from the selected side.

The system has been implemented using Java and the JTreeMap library for coding and MySQL for data storage. The system presents all the data in a couple of seconds and after any interaction with the interface the visualization is refreshed in about two seconds. *Figure 1* shows one of the views of an EIS (Executive Information System) development and implementation project used as an example to evaluate the representation with real phases and projected dates. The view in the figure presents the cost deviation for the design tasks. Some of the measures have been estimated for evaluation purposes since not every measure by phase had been previously recorded and, therefore, no data was available for some of the measures.

4. Experiment overview and results

Eighteen computer science students at the UPV/EHU (University of the Basque Country) participated in the experiment that we designed to examine the efficiency and effectiveness of the tool. The students were randomly assigned and they were reasonably familiar with the notions of projects, tasks and measures. We wanted to compare the results of people performing information retrieval tasks with the Cube system with those of people using another tool to know how well the system would assist these tasks. We were not able to find a similar system containing the same information, and therefore, we considered comparing the T-Cube tool with a set of reports. These reports were specifically designed to find the answers to the questions and they included Gantt diagrams also.

We divided the users into two groups: the first group answered some questions using the T-Cube and the second group used only the reports to answer to the same questions. For the experiment, the second group of users was provided with printed versions of the reports, which were also available through a reporting tool with links to these reports.

We selected for the experiment typical information retrieval tasks of project management, which included plan tracking and deviation, error tracking, change management, deliverable management and some other tasks that required a more complex analysis with multiple filters and measures. These information retrieval tasks were organized into 15 questions that the users had to answer using Moodle platform (a free course management application) using personal computers in both cases. The questions are listed in *Table 2*.

Table 2.- Questions examined
1.- Identify the task with the greatest deal of real effort
2.- Identify the task that has a cost deviation greater than 6000
3.-Identify the task with the maximum number of total errors.
4.-Identify the task with the greatest deal of error correction effort
5.-Identify the task with the maximum number of total errors among the tasks with code error correction effort.
6.-Identify the phase of the project with the highest number of requested modifications.
7.-Identify the task with the highest number of rejected modifications
8.-Identify the type of task with the highest number of rejected deliverable
9.-Compare a set of modules and identify the one that has the maximum number of pending deliverables.
10.-Identify the task of the “requirements analysis” type that has requests for modifications and errors.

The experiment consisted of an only session that each person began by listening to an explanation about the information that was going to be presented. For the first group the next step was learning how to use the visualization tool and for the second group getting familiar with the reports. After some minutes, the participants answered to the questions and the first group completed a subjective questionnaire concerning the tool.

We will now analyze the time to successfully complete tasks using the T-Cube or the reporting tool and printed reports. The values below correspond to the average completion time in seconds (for correct answers only) as function of the tool. Number of correct answers is indicated in parenthesis.

Table 3.- Average completion time

Question	Using report tool	Using Cube
1	193,3 (8)	80,6 (8)
2	130,0 (8)	58,6 (8)
3	139,7 (7)	67,1(8)
4	84,9 (8)	82,3 (8)
5	180,8 (8)	156,4 (7)
6	115,0 (7)	214,9 (8)
7	34,3 (7)	95,6 (7)
8	130,0 (7)	122,0 (8)
9	33,1(8)	64,1 (8)
10	121,2(6)	112,6 (7)

The T-Cube tool seems to be more efficient since in 7 of the 10 questions the completion time with Cube was shorter. The mean value for all the questions is 105.4 seconds for Cube and 116.2 for the reporting tool. Concerning the total number of correct answers, the value is 77 using the Cube versus 74, so we could conclude that the tool is more effective. However, despite the first impression of the use of the new proposal, a t-test on the means of the Time and Score variables, using the Tool (Cube or Paper) as the only independent variable, does not support the initial assessment. The t test (or F test) gives 0.3111 with p=0.760, well above the level of 0.05, which means that the null hypothesis of equal means cannot be rejected. The same happens with the total number of correct answers with t = 0.4291 (p=0.674). Therefore, further research into the aspects that lead us to a good perception of the new method should be carried out.

In the subjective evaluation we presented nine questions, eight of them to evaluate the utility of the tool for particular tasks and a last one to evaluate the utility globally. The values below correspond to the percentage of students that evaluated the tool as a function of the mark (1=not useful at all, 5=very useful)

Table 4- Evaluation of usability per task

	1	**2**	**3**	**4**	**5**
1. Get an overview of costs and efforts of the tasks.	0	12.5	37.5	25	25
2. Find a task with a cost deviation higher than a given value.	0	0	50	25	25
3. Find the phase of the project with the highest number of requested modifications.	0	12.5	25	37.5	25
4. Find the task with the greatest deal of code correction effort and the number of errors associated.	0	12.5	37.5	12.5	37.5

5. Identify the task with the highest number of modifications rejected.	0	0	50	12.5	37.5
6. Compare a set of modules and identify the one with the maximum number of pending deliverables.	0	25	0	50	25
7. Identify the type of task with the highest number of deliverables rejected.	0	0	37.5	37.5	25
8. Make a complex analysis requiring multiple filters and visual analysis.	0	12.5	37.5	37.5	12.5
9. In general, the tool has been helpful to answer the questions.	0	12.5	50.0	0	37.5

In summary, users evaluated the tool quite positively. The results of the final assertion show that 37.5% of the users evaluate the utility of T-Cube as 5, the maximum possible, and 50 % of the rest assess it at 3. The average values seem to indicate that the user liked the tool, since 26.56% of the answers are 5 and 26.69% of them are 4. The questions where the users indicated slightest agreement (questions 4 and 6) are those that include comparing groups that were on different sides of the cube and that required the user to consider information not in the current view.

5. Conclusion

This approach presents a general overview of a software project development process, which is divided into tasks and analyzed from different perspectives or hierarchies. The system offers the possibility of choosing the hierarchy for structuring the information and focusing on the desired item of the hierarchy. Cube presents information about effort and cost deviation, deliverable tracking, error types, effort distribution, and facilitates scope management in a single tool, and the user can analyze one aspect or another by choosing from a list of measures, a more effective alternative than others as extended as Gant diagrams. The wide range of measures presented and the flexibility in the organization of the tasks offers a high number of possibilities to analyze the project from different perspectives.

Users have expressed a positive judgment on the presented visualization, which is also important since the success of an information visualization tool depends partly on users' subjective opinions of the tool's interface and utility.

References

[1] N. Fenton, and S. Pfleeger, *Software metrics: a rigorous and practical approach*: PWS Publishing Co. Boston, MA, USA, 1997.

[2] J. Stasko, “Three-Dimensional Computation Visualization,” *Proceedings on the 1993 IEEE Symposium on Visual Languages*, pp. 100-107, 1993.

[3] M. Baker, and S. Eick, “Space-filling Software Visualization,” *Journal of Visual Languages and Computing,* vol. 6, no. 2, pp. 119-133, 1995.

[4] D. A. Umphress, T. D. Hendrix, I. James H. Cross *et al.*, “Software visualizations for improving and measuring the comprehensibility of source code,” *Sci. Comput. Program.,* vol. 60, no. 2, pp. 121-133, 2006.

[5] C. d. Souza, J. Froehlich, and P. Dourish, “Seeking the source: software source code as a social and technical artifact,” in Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work, Sanibel Island, Florida, USA, 2005, pp. 197-206.

[6] M. Ogawa, and K. Ma, “StarGate: A Unified, Interactive Visualization of Software Projects,” in PacificVIS'08. IEEE Pacific Visualization Symposium, 2008, Kyoto, 2008, pp. 191-198.

[7] N. Fenton, and M. Neil, “Software metrics: successes, failures and new directions,” *The Journal of Systems & Software,* vol. 47, no. 2-3, pp. 149-157, 1999.

[8] F. H. Damborg, and M. Lars, “A contextual approach to improving software metrics practices,” *IEEE Transactions on engineering management,* vol. 55, no. 4, pp. 602-616, 2008.

[9] J. Stott, P. Rodgers, R. Burkhard *et al.*, “Automatic Layout of Project Plans Using a Metro Map Metaphor,” in Proceedings on the Ninth International Conference on Information Visualization, 2005, pp. 203-206.

[10] K. Hansen, “Project Visualization for Software,” *IEEE Software*, pp. 84-92, 2006.

[11] J. Jones, M. Harrold, and J. Stasko, “Visualization for Fault Localization,” in Proceedings of ICSE 2001 Workshop on Software Visualization, Toronto, Ontario, Canada, 2001, pp. 71-75.

[12] P. Zhang, and D. Zhu, “Information Visualization in Project Management and Scheduling,” in Proceedings of the 4th Conference of the International Society for Decision Support Systems (ISDSS'97), 1997, pp. 1-9.

[13] R. Burkhard, and M. Meier, “Tube Map Visualization: Evaluation of a Novel Knowledge Visualization Application for the Transfer of Knowledge in Long-Term Projects,” *Journal of Universal Computer Science,* vol. 11, no. 4, pp. 473-494, 28/04/2005, 2005.

[14] K. El Emam, and A. Koru, “A Replicated Survey of IT Software Project Failures,” *IEEE Software,* vol. 25, no. 5, pp. 84-90, 2008.

[15] P. Bannerman, “Risk and risk management in software projects: A reassessment,” *The Journal of Systems & Software,* vol. 81, no. 12, pp. 2118-2133, 2008.

[16] T. DeMarco, and T. Lister, *Waltzing with Bears: Managing Risk on Software Projects*: Dorset House Publishing Co., Inc. New York, NY, USA, 2003.

[17] B. Boehm, “Software risk management: principles and practices,” *Software, IEEE,* vol. 8, no. 1, pp. 32-41, 1991.

[18] B. Boehm, “Tutorial: Software Risk Management,” *IEEE Computer Society*, 1989.

[19] IEEE, "IEEE Std 982.1-1988 IEEE Standard Dictionary of Measures to Produce Reliable Software," 1988.

[20] IEEE, "IEEE std 1028-1997 IEEE Standard for Software Reviews and Audits.," 1989.

[21] J. Van Wijk, and H. Van de Wetering, "Cushion treemaps: Visualization of hierarchical information." pp. 73-78.

[22] W. Jungmeister, “Adapting treemaps to stock portfolio visualization,” 1992.

[23] L. Jin, and D. Banks, “TennisViewer: a browser for competition trees,” *IEEE Computer Graphics and Applications,* vol. 17, no. 4, pp. 63-65, 1997.

[24] B. Shneiderman, S. Card, and J. Mackinlay, *Readings in Information Visualization: Using Vision to Think*: Morgan Kaufmann, 1999.

[25] M. Eppler, “The Image of Insight: The Use of Visual Metaphors in the Communication of Knowledge,” in Proceedings of I-KNOW '03, Graz,Austria, 2003, pp. 81-88.

Visual Interactive Exploration of Spatio-Temporal Patterns

Radoslaw Rudnicki, Monika Sester, Volker Paelke
Leibniz Universität Hannover, Appelstraße 9a, 30167 Hannover

Abstract

In many applications it is of high interest to analyze spatio-temporal (ST) datasets for recurring patterns. In our approach we examine spatio-temporal datasets for movement patterns using efficient query techniques. In an iterative process combining search and interaction an expert can review search results and either draw direct conclusions and annotate the pattern or refine the search pattern. The key feature of the approach is that patterns are not specified in advance but established from the data set.Futhermore, they can occur in multiple levels of detail, thus are designed to be scale invariant.

1. Introduction

The search for spatio-temporal patterns in large data sets becomes an increasingly relevant task in many applications. Often the patterns of interest are not known in advance or not well defined. To address this problem where an a-priori definition of patterns is not feasible or useful, we propose an experimental approach in which a human expert marks patterns of potential interest in a visualization of "temporal aggregates" of the raw data. Similar patterns can then be identified automatically by searching the dataset for similar occurrences. After a search the expert can review the results and either draw direct conclusions and annotate the pattern or refine the search pattern (if the results are not yet sufficient).

For the analysis of patterns it is first of all necessary to define "patterns". The central characteristic of a pattern in our approach is that the same spatio-temporal formation of data elements occurs multiple times. However, not every coherent spatio-temporal data element is a relevant pattern. We name reoccurring patterns "proto-patterns"; the importance of these is apriori unknown. E.g. on a straight street segment many sequences of straight movement could be identified as "proto-patterns" but this might not be as interesting for an application as a single criss cross trajectory in the same area. For our approach we define patterns as elementary parts and/or aggregates of trajectories on a grid. Relevant patterns correspond to interesing constellations among these trajectories, where interesting is defined as an information theoretic measure, namely the frequency of occurrence. For each pattern a Fourier Descriptor is calculated to describe the shape of the patterns.

2. Related work

Closely related to our work are approaches that simplify and generalize patterns using methods from geometric generalization, e.g. [1]. The use of databases for spatio-temporal processing was recommended for example in [2]. P. Laube et al., worked on discovering defined patterns, e.g. the Flock and Encounter-Pattern in [3].

In [4] Dodge, Weibel & Lautenschütz proposed a taxonomy to classify and categorize movement patterns. In order to determine the similarity of geometric features, appropriate measures have to be devised. Popular methods are geometric moments as proposed by Heinzle et al. in [5]. Hild used affine invariant moments and Fourier Descriptors, to describe geometric features that are used as ground control features for automatic image registration in [6]. In [7] Lee et al. measured similarity of shapes also with fourier discriptors. B. Srinivasa Reddy and B. N. Chatterji describe in [8] a technique for translation, rotation, and scale-invariant image registration. Andrienko & Andrienko use aggregation of ST patterns mainly to enhance visual inspection in [9].

3. Describing and indexing ST-patterns

The base information in our application are trajectories. Trajectories consist of sequences of coordinates with timestamps. Timestamped data often appears in irregular intervals and time can be desynchronised between different observed objects. Therefore the first step is to resample and align the data. The goal is to resample the trajectories without changing their characteristics.

For our use case, linear interpolation is sufficient, because minor changes in the coordinates of a point (at the scale of a few centimeters) do not have a significant effect on the resulting patterns and are typically below the threshold of the accuracy of the sensors used in the acquisition of the trajectories.

Synchronized datasets simplify the analysis of a situation at an arbitrary point in time. The human operator has to adjust the resampling rate depending on the spatial and temporal extent of the patterns of interest, taking the sampling theorem into account.

From the original trajectory data, which only includes the positions and timestamps, other information of the observed objects like direction and speed can be derived at any time step. This allows to aggregate the data into cells and simplifies the following operations and calculations by reducing the size of large datasets to manageable proportions. The aggregation removes details and helps to focus on the essential characteristics. It is obvious that the selection of an appropriate cell size by the user is essential to achieve a suitable balance between data reduction and the potential loss of information.

The smaller the cell size, the less the spatial data is aggregated and the data cells provide a closer approximate of the original trajectory data. The bigger the cell size, the more trajectories are integrated, which reduces the amount of data but also brings about the possible loss of essential information so that patters that were part of the original dataset can no longer be detected. On the other hand, it leads to a generalization of the spatial situation, allowing an efficient overview of the whole scene or trajectories.

The same is true for temporal aggregation: Using very short time intervals all patterns are nearly identical, because almost all movement is a short step forward. Thus, there is no chance to build complex patterns. In very long intervals the data becomes more individual, and it is hard to find any general structure in it. So, the best option is somewhere in between, depending on data and also on the application.

For each cell all crossing trajectories are considered and suitable aggregates of speed, direction and density information within a given interval are calculated and stored as attributes of the cells.

An aggregation of neighboring filled cells defines proto-patterns. The boundary of these aggregations describe the spatial contour of the pattern. To identify similar patterns, it is important to define a measure of similarity. We examined two approaches to similarity analysis. The first approach just uses geometric properties of the data, and ignores the attributes speed, density and direction, i.e. the cell value is set to true, if there is some trajectory information inside and false if not. It is very fast but can't address all cases of similarity and dissimilarity that are relevant. The second approach is mor precise as it takes all available data into account.

For comparing such patterns it is useful to define similarity measures that are independent of scale, specific location and orientation. In the first approach, we experimented with the one dimensional discrete fourier transformation (1D DFT (see e.g. Lee et al. [7]). In order to achieve invariance in shape, size and positioning all objects contours are resized to a specific length and shifted to the point of origin. By resampling the outline and interpreting the x and y values as real and imaginary part of a complex number, it is possible to determine a vector which represents this shape. After normalising the vectors the euclidian distance represents a similarity measure between shapes.

This method produces results that comply with human assessment of similarity if the cell size is small enough and the speed, direction and density values can vary without having an influence on the importance of patterns. The problem with the cell size is that a rotated shape can optically depart from the original one, because the impact of discretization becomes too strong. Therefore it is useful to smooth the contours to keep pattern rotation invariant. Figure 1 illustrates the problem with a simple shape (1.), which is rotated about 45° (2.). It looses a lot in contour similarity. (3.) shows that smoothing contours reestablishes similarity.

Another problem with this approach is, that not all shapes can be distinguished reliably. Because of sampling only the contour, holes cannot be detected. The algorithms can not distinguish shapes, if their outer contour is equal. And there can be some deviations in similarity depending on the starting position for sampling the contour and the sampling rate. Of course, the sampling rate should be at least as twice as high as the highest frequency in this shape (sampling theorem). But the higher the rate, the longer the algorithm takes. With a relatively low rate, the corners can not be represented correctly.

However, as mentioned above, this first approach does not take all requirements and options into account, as the aggregated trajectory attribtes speed, direction and density are not used. It is like a 2D-region of *true* and *false* values from which the contour is extracted. Therefore we extend the 2D-regions from a binary representation to higher dimensional

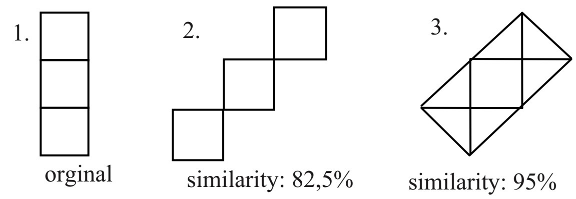


Figure 1. ***Comparison of aggregated and discretisized trajectories withand without smoothing***

region including more information, namely density, speed and direction. This was implemented in a second approach. The central idea is to use the different features as multiple channels of overlapping information, similar to color channels in a multicolor picture. A pixel is represented by a cell, including different data channels. In our case we have the speed, direction and density as "colors" of the picture instead of RGB. With this special view on the data, a 2D discrete fourier transformation (DFT) can be used. A mere 2D DFT on a picture is not invariant to rotation, but this can be solved by an additional intermediate step in an extension of the 2D DFT known as the Fourier Mellin transformation [11], see Figure 2.

The first row shows two original data sets, a horizontal and a vertical line. In the second row is their logarithmic visualization of the frequency space after DFT. Due to the different orientations, the visualization of the frequencies also is rotated. The third row shows the result of the Melling transformation: nearly the identical image, indicating the rotation invariance achieved by representing data in polar coordinates.

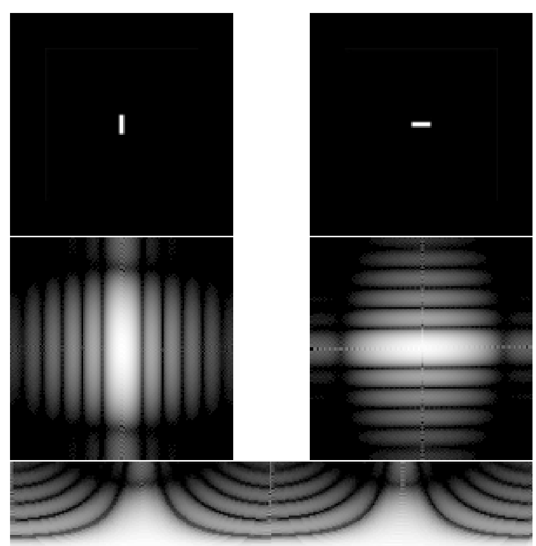


Figure 2. ***Steps of a Fourier Mellin Transformation***

The second approach includes all requirements: the similarity measure is invariant to changes in scale, position and orientation, and all characteristics of speed, direction and density are taken into account. A key improvement is that a distinction of patterns that are of similar shape in their spatial aggregates can be made between those with and without holes. If the cell size is small enough this approach produces promising results. For bigger cells, there is still the need to smooth the contours. These representations are used to calculate the differences between different proto-patterns. Hereby the closest matches of a pattern can be found.

4. Implementation

The resampling of the trajectories is straightforward. Assume that we have a starting point p_i, two positions p_n and p_{n+1} with timestamps t_n and t_{n+1} and the next synchronized timestep t_a is between these both, a new synchronized point is created by linear interpolation:

$$p_{new} = p_i + t_a \cdot (p_{n+1} - p_n), n \geq 0, 0 \leq t_a \leq 1.$$
$$p_i = p_n$$

For an efficent implementation of resampling, it is useful to have quick access to data belonging to a certain trajectory at a requested timestamp. Because datasets are typically large, an out-of-core method is needed which allows to work on it without keeping it all in RAM. Therefore a MySQL database is used. It allows to use index structures to read data, saves the data efficently and may use more space than RAM allows. So, a table *trajectory* with attributes *id*, *position_x*, *position_y*, *timestamp* with the *id* as primary key and a BTree index on *timestamp, position_x and position_y* is created. A table *trajectory_sampled* with the same attributes and indexes holds the resampled data also efficient accessable.

Grid cells intersection points with the trajectories were calculated. In a further table *grid* with attributes *x*, *y*, *time_step*, *trajectory_id, speed* and *direction* the intersected cells are saved including the time step information. *x* and *y* do not represent the position itself, but the grid cell. Indexes are used on *x*, *y* and *time_step, which* accelerates the next steps in data processing.

Knowing when which grid cell is intersected by which trajectory, enables the aggregation inside the cells. Therefore the speed and direction information of different trajectories inside a cell are averaged. To average the direction, the vector of movement during the considered time interval is transformed to polar coordinates and reduced to its angle α. The average of the speed attribute is straigthforward.

The result of this last step are saved in a table *aggregated_data*. As additional new information, the density of trajectories inside a cell is added. So we have the attributes *x* and *y* representing the cell, *angle*, *speed*, *density* and *time_step*.

The identification and delineation of coherent shapes from the individual grid cells is performed in a region growing fashion. To this end, randomlyan intersected grid cell is taken from the table *grid* and examined with respect to its neighbours (N8-neighborhood) which are also marked as intersected at this moment or time interval, respectively. All localised new cells again examine their neighbours, recursivley. In this way the composition of cells, that form a pattern can be found. This is repeated until all intersected cells from the database are allocated to a proto-pattern. Because of the indexes on the attributes *x, y* and *time_step* the neighbours can be checked very efficently. With the data from aggregated_data a new table pattern (pattern_id, x, y, speed, angle, density, time_step) can be created. It is the first table, that uses aggregated data and saves, which cells belong to which proto-patterns at which time step.

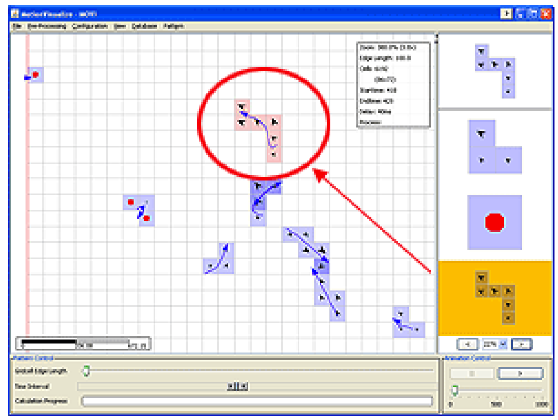


Figure 3. ***Screenshot of the software prototype for visualization and interaction***

The concept is implemented as a prototype written in Java. The time interval and the grid size can be easily controlled by the user with sliders. To visualize the different components of a cell, the density of the trajectories are visualized using the saturation of their background color. The direction of the aggregated data is shown by an arrow. The size of this arrow shows the average speed inside this cell. The software also allows to have a look at all proto-patterns and by clicking them to jump to the time step in which they appeared. After chosing a proto-pattern there is the option in the software to find most similar patterns.

The screenshot in Figure 3 shows the GUI of the software prototype. The main panel is subdevided in cells. The right side informs about proto-patterns. It lists all available proto-patterns which were found during processing and allows the user to select them for further analysis. The user can compare the selected pattern with other patterns and order them by similarity. The similarity measure gives a ranking from very similar to not similar, however, it does not define a precise cutoff value for similarity. Therefore, the user can identify the first pattern in the ordered list, which he consideres no longer to be similar and thus determine a threshold. After having defined this similarity measure, the numbers of instances of the different patterns can be counted. The number of similar proto-patterns shows, if this pattern appears many times or is an exception.

5. Observations and Discussion

The use of the database and especially its index structure seems to be a good backend of the implementation. Besides the efficient disk space use and quick return of requested data, more efficient use of multi-core processors is made. While one core works on processing the data, another core is adapted for reading and writing the database. Also workload of RAM can be reduced. Alternatively it would be possible to create a custom index structures to solve this problem.

In the two approaches, different ways of using the frequency space were described. The first approach works with the contour of a shape. The second idea, to consider the data matrix as picture is more complex and needs additional procedures to keep rotation invariant, but can differ proto-patterns more precise as Figure 4 shows. The first row of this figure shows the different behavior of the first and second approach, by analyzing shapes with and without holes: in approach 1 the filled shape is considered as the same as the shape with hole, whereas in approach 2 (right hand side) a difference is notified. The second row demonstrates how the algorithms evaluate same shapes with different aggregated information (e.g. with respect to direction). Again the first approach does not react on this significant difference.

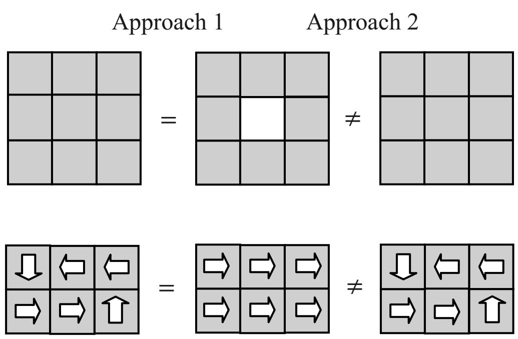


Figure 4. Advantage of Fourier Mellin Transformation compared to 1D Fourier Transformation

One important observation when comparing patterns with Fourier Descriptors and differences in their representive vectors of the coefficents is that there is no fixed value which can be used as threshold for similarity. In some datasets a value of 87% seems to indicate high similarity, in other it does not. The range of values isn't fully exploited because, there is always any kind of similarity. In our special case the smallest common shape is a box which can be found in every pattern. Typically we had similarities between 75% and 100% in our dataset.

This is why in our approach the human operator is in the loop to identify similarities of patterns. Based on this, the corresponding similarity values can be calculated and used as thresholds in subsequent processes.

6. Summary and Outlook

The basic idea of our approach is to identify "interesting" patterns in trajectories and store and index them in a database. From the frequency of occurrence of a pattern application dependent inferences can be drawn: when looking for dominant patterns, the ones with a high frequency of occurrence are chosen; when looking for non-conform patterns, e.g. unusal movement patterns in a large crowd, then the most seldom patterns are important.

It remains to be seen, if the ranking of all patterns agree with human evaluation in detail. Therefore user tests are necessary, which will be conducted in the near future.

7. References

[1] H. J. M. L. Yiu, X. Zhou, C. S. Jensen, H. T. Shen, (2008): Discovery of convoys in trajectory databases, in: Proceedings of the VLDB Endowment Volume 1 , Issue 1

[2] Ouri Wolfson, Prasad Sistla, Bo Xu, Jutai Zhou, Sam Chamberlain, (1999): DOMINO: Databases fOr MovINg Objects tracking, ACM SIGMOD Record Volume 28, 547 – 549

[3] Laube, P., Imfeld, S., & Weibel, R., (2005): Discovering relative motion patterns in groups of moving point objects, International Journal of Geographical Information Science, 19(6), 639–668.

[4] Dodge, S., R. Weibel & A.-K. Lautenschütz, (2008): Towards a taxonomy of movement patterns. Information Visualization, 7, 240-252.

[5] F. Heinzle, K.-H. Anders and M. Sester, (2006): Pattern Recognition in Road Networks on the Example of Circular Road Detection, Geographic Information Science , no. 4197 , p. 253-267 , Münster

[6] Hild, H., (2001): Automatic Image-To-Map-Registration of Remote Sensing Data, in: D. Fritsch & R. Spiller, eds, 'Photogrammetric Week '01', Herbert Wichmann Verlag, Heidelberg, pp. 13–23.

[7] D.J. Lee, S.Antani and L. R. Long, (2004): Similarity Measurement Using Polygon Curve Representation and Fourier Descriptors for Shape-based Vertebral Image Retrieval

[8] B. Srinivasa Reddy and B. N. Chatterji, (1996): An FFT-Based Technique for Translation, Rotation, and Scale-Invariant Image Registration, IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 5, No. 8

[9] Andrienko, G., & Andrienko, N. (2008): Spatio-temporal aggregation for visual analysis of movements, In Proceedings of IEEE Symposium on Visual Analytics Science and Technology (VAST 2008), IEEE Computer Society Press, pp.51-58.

On the Usability of Reverse Engineering Tools

F. Ferrucci, R. Oliveto, G. Tortora, G. Vitiello
DMI – University of Salerno
84084, Fisciano (SA), Italy
{fferrucci, roliveto, tortora, giuvit}@unisa.it

S. Di Martino
University of Naples "Federico II"
80126, Naples, Italy
dimartino@na.infn.it

***Abstract* – Usability guidelines and evaluation methods for Reverse Engineering and Re-engineering tools are still missing. This paper presents the current state of the art, carefully analyzing and searching for case studies, structured demonstrations and other empirical work from which some design best practices could be deduced in the form of requirements, guidelines, heuristics, or maxims for Reverse Engineering and Re-engineering tools. The derived set has been analyzed to merge the similar items. As a result, a list of usability guidelines/requirements for the design of Reverse Engineering and Re-engineering tools is presented.**

I. Introduction

Usability is a key quality factor of Reverse Engineering and Re-engineering tools and the success of these tools critically depends on it. As a matter of fact, usability is often the major concern in the adoption of software engineering tools in industry. If a tool is difficult to use, it will hardly be adopted, no matter how useful it may be. Analogously, if a tool provides information in a form that does not support user cognitive process, it will not be effective and thus it will be discarded.

In the context of software systems, a wide body of knowledge has been built about usability and quality of use, which includes both usability guidelines and usability evaluation techniques. Such guidelines are useful for the design of user interfaces and also support expert evaluators during the usability assessment process. Usability evaluation methods have been developed in order to assess the efficiency, interaction flexibility, interaction robustness, and quality of use of user interfaces. Usability evaluation techniques can be broadly divided in two groups: expert evaluation and user studies. Expert evaluation techniques include cognitive walkthrough, model-based evaluation, heuristic evaluation, and are used since the initial phases of the software development cycle. They allow discovering usability problems early during the design phase and avoid waste of resources, before a working prototype is available. On the other hand, user studies are based on the direct involvement of users through observations, interviews, and questionnaires and include empirical methods and observational methods. This kind of study can be useful not only to assess usability of a specific system but also to derive, based on empirical evidence, guidelines that can be applied more generally.

However, as for Reverse Engineering and Re-engineering tools, standard usability guidelines and evaluation methods are still missing. In fact, these tools are often based on abstract tasks such as "understanding program" or "implementing changes based on the supplied information" that make usability characterization and evaluation quite hard. In addition, interfaces for these tools include a set of 2D (visual diagrammatic languages) and 3D structures that can be characterized by several different properties such as layout algorithms, layered view, fisheye view, navigation 3D, which may have a great impact on the tool usability (possibly causing excessive cognitive workload, lost in space, etc.).

Yet, usability issue for Reverse Engineering and Re-engineering tools is considered a crucial quality factor and several researchers have tried to face this problem. Some researchers, based on their experience with the design and use of these tools, have suggested some specific requirements and factors that should be considered when developing such tools. Others have reported on the analysis of a specific system, from which some hints can be also drawn. However, so far no comprehensive list of these suggestions can be found in the literature. Thus our goal has been to study the current state of the art carefully analyzing and searching for case studies, structured demonstrations and other empirical work from which some design best practices were deduced and presented in the form of requirements, guidelines, heuristics, or maxims. Each paper has been examined, in order to extract knowledge about Reverse Engineering or Re-engineering requirements and guidelines. This initial list presented a high level of redundancy. Thus, we have merged the sentences presenting overlap, obtaining a final list of guidelines/requirements for the design of Reverse Engineering and Re-engineering tools. The list contains both general usability guidelines that may fit any Reverse Engineering and/or Re-engineering tool, and specific guidelines, suited for some specific tasks composing the Reverse Engineering and/or Re-engineering processes.

The paper is organized as follows. In Section 2 we recall the main concepts related to Reverse and Re-engineering techniques, while Section 3 presents some international standards for software product quality and usability. In Section 4 we provide the set of usability guidelines/requirements for designing Reverse and Re-engineering tools, which result from the analysis we have performed on the existing literature. Finally Section 5 gives final remarks and discusses directions for future work.

II. Reverse Engineering And Program Comprehension

Reverse Engineering is defined as "the process of analyzing a subject system to (i) identify the system's components and their interrelationships; and (ii) create representations of the system in another form or at a higher level of abstraction." [5]

However, this definition presents some limitations [26] since there are several Reverse Engineering tools that do not produce as output high level components and relationships (e.g., slicing). Moreover, in the above definition, it is not clear the meaning of "representations of the system in another form". Thus, Reverse Engineering can be considered as "every method aimed at recovering knowledge about an existing software system in support to the execution of a software engineering task." [26]Starting from a software product, a Reverse Engineering process should allow deriving design decisions even when little or no documentation about the product is available. The techniques adopted in Reverse Engineering are commonly used to support software engineers in software maintenance activities, such as the exploration of a software system to determine where modifications should be done. In fact, maintenance activities may only succeed when the software system is adequately understood. To this aim Reverse Engineering techniques and tools allow software engineers to build a descriptive view of the system at various levels of abstraction, with the goal to convey as much knowledge as possible about the software product and about the critical design choices done by developers. Reverse engineering are "human intensive" activities, especially when performing program comprehension tasks [26].

Methods and tools to support Reverse Engineering vary deeply depending on the specific tasks required and on the given context, but also depending on the overall target of the Reverse Engineering activities performed. Two different approaches are employed, depending on whether *static* or *dynamic* analysis is performed. Static analysis handles source code, structure, architecture and documentation and derives knowledge about the system or component without executing the program [7]. Differently, dynamic analysis activities require running the system only under controlled circumstances, where results are often known a priori, with the goal to derive run-time information and understand dynamic features of a design, such as execution sequences and relative time order of events [27]. Compared with static analysis techniques, the causes of bugs and system irregularities may be most easily found through dynamic approaches, especially in legacy systems that are process centric rather than data centric. However, in spite of their high potentiality in providing program understanding, dynamic analysis techniques are adopted less frequently than static analysis. This is mainly caused by the higher difficulties that raise both in the collection of large amounts of information during execution, requiring special settings in the environment, and in the interpretation of this information to extract useful knowledge about the system.

A. Program Comprehension

The definitions of Reverse Engineering reported in the previous section have highlighted the stronger relationship between Reverse Engineering and software comprehension. Indeed, *Program Comprehension* is meant to complete the Reverse Engineering process and is a major factor in providing effective software maintenance and enabling successful evolution of computer systems. After a program is either statically or dynamically analyzed, software engineers are supposed to spend a substantial amount of time in understanding the gathered information, before performing adaptive, corrective or perfective maintenance tasks or reuse or code leverage tasks. These activities benefit of abstract representations of a software system that are usually conceived to capture some features of the system in a form that better supports the target Reverse Engineering tasks, with respect to the collected low level code information [2].

1) Human Cognitive Models

To accomplish program comprehension tasks, programmers employ various strategies and use cues in code or documentation (when available) as guidance. The construction of several complementary views, able to capture different features of a software system, is a common approach for a comprehensive understanding of the system. Exploring how human mental processes work can help us understand program comprehension better. The process of how humans understand code has been extensively investigated [27]. Often software engineers use programming knowledge, domain knowledge and various strategies to understand a new piece of program [26]. For example one might rely solely on source code to extract syntactic abstractions. Researchers have proposed several models for program comprehension and have concluded that software engineers will try to construct mappings through *mental models* during a program understanding process [27]. A mental model is defined as "an internal, working representation of the software under consideration" [27].

Cognitive models are often used to represent the mental processes software engineers use and the interactions that happen between humans and machines. They correspond to the mental models used by software engineers as they form a mental representation of the program under study. In research about how human beings acquire knowledge, several cognitive models have been proposed [18]. It is clear that the more a tool is able to support the user's cognitive model the more effective it is.

2) Software Visualization

Software visualisation is the process of modelling software systems for comprehension [16]. It consists in the application of graphical techniques to represent different aspects of software – such as the source code, the software structure, runtime behaviour, component interaction, or software evolution – in order to reveal patterns and behaviours that inform software comprehension through all stages of software development [15]. Good overviews of traditional visualization techniques can be found in [6], [20].

Researchers agree that the appropriateness of a given technique is tightly related to the target task [16]. Several taxonomies have been proposed to establish crucial attributes of a good software visualization tool [12], [29] [22], [15].

The classical taxonomy by Price, Small and Baecker suggests six major categories of attributes: *scope* (the range of programs the system can take as input), *content* (the subset of information about the software that is visualised by the system), *form* (the characteristics of the output of the visualization), *method* (how the visualization is specified), *interaction* (how the user interacts with and controls the visualization), and *effectiveness* (how well the system communicates information to the user) [17]. Other taxonomies are provided in [29], [15].

III. Usability In Reverse Engineering Tools

In the present section we analyze some international standards provided in the fields of software engineering and information technology that address usability as a key factor for software product quality.

Other than traditional standards for usability, such as the ISO/IEC-9126 (2001), or the ISO/IEC 14102, a good method for the usability evaluation of Reverse Engineering and reengineering tools is suggested by the International Standard ISO/IEC 25062:2006 - *Software engineering — Software product Quality Requirements and Evaluation (SQuaRE).* It includes a standard method for reporting usability test findings, the *Common Industry Format* (*CIF*), that could be profitably adopted for the evaluation of software maintenance methods and tools. The CIF is appropriate whenever quantitative measurements are collected as a result of a formal usability test and is particularly suitable for summative/comparative testing. This format does not specify the method to perform a usability test but it is rather meant to provide guidance on how the results of the test should be reported in order to facilitate subsequent analysis. It includes eight key elements:

- the description of the product,
- the goals of the test,
- the test participants
- the tasks the users were asked to perform,
- the experimental design of the test,
- the method or process by which the test was conducted,
- the usability measures and data collection methods, and
- the numerical results.

The concept of usability expressed in terms effectiveness, efficiency and subjective satisfaction can be adapted to the reverse engineering and re-engineering fields. In particular, according to Systä and Väänänen-Vainio-Mattila "Effectiveness refers to the amount of problems that can be solved by using the specific tool and efficiency implies the amount of effort needed to solve the given reverse engineering or re-engineering tasks. Subjective satisfaction is related to the professional users' satisfaction of using the tools in their daily work over a long period of time. In addition, the organizational satisfaction with the tools becomes a broader context for tool satisfaction" [23]. Underlying all this factors, scalability is an important specific issue to be considered for the usability of reverse engineering and re-engineering tools, as it largely impacts on the effectiveness, efficiency, and satisfaction of the tool when applied to real size industrial software.

Another important aspect to highlight is the nature task-oriented and user-oriented that reverse engineering tools should satisfy. Indeed, the information recovered from the code can be useful or less depending on the Reverse Engineering task to perform and on the skill of the software engineer and on the amount of knowledge he/she already has about the system [26].

IV. Usability Guidelines and Requirements

In this section we present the set of usability guidelines and requirements for designing Reverse Engineering and Re-engineering tools that we have derived from the literature. To develop the following list, we started from the work of Mealy *et al*. [13] that addressed the same issue for the software refactoring tools. Indeed, they reviewed 11 collections of usability guidelines for software systems and combined these into a single list of 29 guidelines. To this list they added other six guidelines based upon Fitts' List for task allocation, and literature from the area of situation awareness and automation. Adopting a similar methodology we expanded this list by adding further 65 guidelines/requirements that are suitable for Reverse Engineering and Re-engineering tools. We carefully searched the literature, looking for journal and conference papers dealing with experiences with tools of Reverse Engineering and Re-engineering, e.g. case studies, structured demonstrations, cognitive studies.

Each paper was examined, looking for requirements, heuristics, guidelines, rules, heuristics, and maxims for Reverse Engineering and Re-engineering tools. This knowledge was extracted to obtain an initial list that obviously presented a high level of redundancy. Then, the authors merged all the sentences presenting a high degree of overlap, and finally all the guidelines and requirements were organized in (i) generic guidelines for Reverse Engineering and Re-Engineering tools and (ii) guidelines for specific tasks. In the following there is the full list of the distilled and additional usability guidelines.

A. Generic Guidelines

Reverse Engineering environment should provide the following basic features:

- the tool needs to support broad search capabilities, including basic, wildcard, customizable searches, with search history tracking [19] [7];
- visualization and searching facilities should complement each other [1];
- the tool should support multi-users [1] [9];
- a history of browsed locations should be provided for all the elements of a Reverse Engineering tool (e.g. reports) to help going back to interesting parts already shown [1];
- tools should provide documentation capability [1];

- tools should provide printing capabilities to show the results of the Reverse Engineering tasks [1];
- the visualization should achieve high completeness and fidelity and support the representation of anomalies [9];
- the visualization should be structured and include features to aid the user in navigating the visualization, by using, for instance, landmarks to reduce the user's chance of becoming "lost" [3] [9];
- the tool needs to support workspace and session persistence [7];
- the tools should support different platforms [1];
- the interfaces should be minimal, simple to understand, organized, without redundancy, socially relevant and aesthetically pleasing. [13] [18] [14]
- users should browse the visualization by following concepts [9].

1) Flexibility and scalability

Flexibility and scalability are key requirements for a successful Reverse Engineering environment:

- the approach must be flexible so that the results can be applied to many Reverse Engineering scenarios as well as different target domains [25];
- tools require flexible conceptual models that can be modified as the user or task requires [19];
- if users can tailor the new system to fit into their existing environment and work the way they want it to - not the way the designer thought they might want it to work - then the system has a much better chance of success. [25]
- tools should be extensible in parser, user interface, and functionality. For constructing extensible integrated applications from a set of tools there are two basic approaches: tool integration and tool composition [1]. The user would then be able to alternate and select among these tools as required [25];
- the tool needs to be able to handle traces with millions of lines of data easily and effectively [7].

2) Tool Interoperability

- Tools should provide export capabilities, useful to generate other results with the created representation of the source code [1];
- tools must include a standard interchange format or a API to be interoperable [19] and should support integration with other information sources [29];
- the tool needs to be extensible to support different sources and destinations. [7]
- the tool needs to be integrated easily with an IDE and other tools to increase its adoption [7].

3) Consistency

- Ensure that things that look the same act the same and things that look different act different [3] [8] [11] [13] [14] [18] [29];
- be consistent with any interface standards for the domain/environment [8] [11] [13] [14] [18] [29];

4) Errors

- Assist the user to prevent errors (through feedback, constrained interface, user of redundancy) [13] [14] and be tolerant of errors [13] [14];
- provide understandable, polite meaningful, informative error messages [1] [10] [11] [13];
- provide a strategy to recover from errors [11] [13] [14];
- permit reversal of actions/ability to restart [13] [14] [18];
- allow the user to finish their entry/action before requiring errors to be fixed. Do not interrupt the task being completed [13] [18];
- automate error-prone tasks/sub-tasks [13];

5) Information Processing

- Clearly define the real tasks the tool can support in the industrial setting [30];
- Reduce the overall cognitive load, i.e. the amount of attention to accomplish a task [18] [29] [30];
- minimize the number of artifacts that have to be kept in short-term memory, by making the required artifacts easily available to the user, facilitating meaningful encoding, by reducing the uncertainty during exploration and by adopting selection rather than entry, names and not numbers, predictable behavior and access to required data at decision points [13] [29] [30];
- minimize short-term memory fading, by minimizing the time that artifacts have to be retained in short-term memory, and the complexity of tasks between successive artifact acquisitions [30];
- make commands and system responses self-explanatory [4] [13] [14];
- use abstraction or layered approaches to assist understanding [13] [15];
- the tool needs to support various levels of detail, such as high level abstraction, intermediate level abstraction and source code navigation [3] [9] [7] [29];
- provide help and documentation, including tutorials and diagnostic tools [1] [13] [14]
- assist the user to maintain a mental model of the structure of the application system/data/task [3] [13];
- maximize the user's understanding of the application system/data/task at the required levels of detail [3] [13]
- cognitive resources or cognitive processing that are "in the head" of the user can be moved outside, such as annotations onto a notepad [9] [28];
- substitute perceptual operators wherever possible. Tools should support cognition by means of fast operations for human thinking (e.g. edge detection) rather than slower ones [28];
- make concrete the ends-means mapping structures. Tools should provide a display-based problem solving mechanism, assuming as interaction strategy the repeated examination of the display for actions that will progress towards his/her goals [28];
- small changes of content or shifts in attention should not cause major differences in the visualization [3] [29];

- visualization should present as much information as possible without overwhelming the user [3] [29];
- the tool needs to support user orientation, such as the position in the global view (i.e., landmarks) [7].

6) Design for the User

- Characterize the user and match the system to the user (user centered design) [13] [7] [14] [18];
- avoid codes by using user's language [8] [10] [13];
- automate mundane/computable tasks/sub-tasks [13];
- assist the user to understand the tool [14] [18].

7) User Control

- Adapt to the user's ability, allow experienced users to use shortcuts/personalize the system, and use multiple entry formats or styles [10] [13] [18];
- put the user in control of the system, ensure that they feel in control and can achieve what they want to achieve. Allow users to control level of detail, error messages and the choice of system style [4] [11] [13]

8) Goal Assessment

- Ensure the user always knows what is happening. Respond quickly, meaningfully, informatively, and consistently to user requests and actions [8] [11] [13] [14] [18];
- use the fewest number of steps/screens/actions to achieve the user's goal [7] [13] [14] [18] [29] and guide the user to find out what to do next. [8] [11] [13] [14] [18];
- in multiple windows, emphasize the relationship of the open windows to the corresponding composite nodes [21];
- the appearance of the overview window should differ from the general windows (by using, for instance, different background colors) [21].

B. Specific guidelines for Reverse Engineering tasks

This section discusses the guidelines defined for Reverse Engineering tasks.

1) Visualization

- Layout algorithms should be employed to make a graph more readable [1] [21];
- the view should be editable by moving and deleting entities that are currently not of interest [1];
- the view should be layered. Indeed, if a grouping of entities can be done recursively, a layered hierarchy should be generated. Such a hierarchy can be shown either in one window or each layer of the hierarchy is displayed in a separate window [1];
- a fisheye view helps to make a graph with a fixed size more readable by zooming in or out parts of the graph similar to a fisheye lens [1];
- careful consideration must be given to when fine-grained source code representations should be used. The benefits of such an analysis need to be commensurate with the costs [19];
- A good level of automation is required in order to make the visualizations of any practical worth [3];
- the visualization should be linked to the original information it represents (the source code) [3];
- the visualization should support comprehension, construction, evaluation, comparison, and evolution of a multitude of software architectures [9], as well as the representation of dynamically changing software architecture [9];
- the visualization should support the communication of the architecture to intended stakeholders [9];
- users should search or query-drill for arbitrary architectural information; [9]
- the visualization should display a representation of the viewpoint definition [9] and support an appropriate set of dynamic data sources [9];
- the visualization should support association of dynamic events with elements of the software architecture [9] and allow live, noninvasive collection of dynamic data recording it for subsequent replay [9];
- the visualization should support the appropriate types of static software architecture data sources [9];
- the visualization should support the recovery of architectural information from sources that are not directly architectural [9] and accommodate large amounts of architectural data [9];
- the tool needs to support the global view of the trace being explored and various views of data slices from different perspectives [7];
- the tool needs to support filtering of events, highlighting events that have specific attributes [7].

2) Automatic Language Conversion

- Select an economically efficient level of automation beforehand and then improve the generated code manually [24];
- Beyond a certain level, extra percent of automation may blow-up the complexity and the price of the tool [24];
- Determine whether the target and the source languages differ considerably in their basic paradigm. [24]
- The attainable level of automation depends on the amount of interaction with the user [24].

3) Parsing

- The tools must include a standard interchange format or a API to support code parsing [19];
- to avoid redundant work in building and adapting parsers, it would be beneficial to use a common set of robust and correct parsers [19];
- incremental parsing should be supported for systems that are changing during the Reverse Engineering phase. It provides the ability to parse only the parts that changed, thereby reducing the parse time [1]. However, sometimes it is important to reparse the whole source code. [1]
- the parser should be fault tolerant aiming at providing the ability to parse incomplete, syntactically incorrect and uncompilable source code [1];
- the *define* and *undefine* preprocessor commands are important for parsing source code and should be supported by the parser [1];

- in addition to macros, compiler switches – such as include directories – should also be supported [1];
- parse abortable, parse errors viewable during parsing, parse continued at error can help reduce the time to parse the source code without errors to generate the representation in the Reverse Engineering tool [1];
- the parsing results represent the completeness, correctness and type of information of the representation of the source code in the Reverse Engineering tool [1].

V. CONCLUSIONS

The quality of Reverse Engineering and Re-Engineering tools is heavily influenced by their usability, but to date, in the literature there is no comprehensive list of guidelines on how to design them to meet this quality factor. To address this issue, in this paper we have presented a preliminary list of usability guidelines and requirements for Reverse Engineering and Re-Engineering tools. We have devised this list by analyzing and searching for case studies, structured demonstrations and other empirical work presented in the literature on Reverse Engineering and Re-Engineering tools. The guidelines have been structured in two sets: the former containing generic guidelines, and the latter containing guidelines for some specific tasks composing the Reverse Engineering process.

It is worth stressing that the devised list is preliminary and some work is planned as future work. In particular, we foresee to further analyze the literature, looking for other researches conducted on Reverse Engineering and Re-Engineering tools and providing suggestions that can be included in the list. Moreover, we also plan to conduct usability evaluation of some Reverse Engineering and Re-Engineering tools with the aim to assess the effectiveness and the applicability of these guidelines. Finally, empirical studies to identify further guidelines and requirements for these tools are also part of the agenda of future work.

ACKNOWLEDGEMENT

The work described in this paper is partially supported by Italian MIUR project “Integrating automated reasoning in model checking: towards push-button formal verification of large scale and infinite state systems”, grant PRIN-2007-20079E5KM8, and “Fondo ex-60%”, Prof. G. Tortora.

REFERENCES

[1] B. Bellay, H. Gall, An Evaluation of Reverse Engineering Tool Capabilities, *Journal of Software Maintenance: Research and Practice*, vol. 10, n. 5, pp. 305 – 331, 1997.

[2] L. Y. Briand, The Experimental Paradigm in Reverse Engineering: Role, Challenges, and Limitations, in Procs. 13th Working Conference on Reverse Engineering (WCRE 2006), pp. 3-8, 2006.

[3] S. M. Charters, N. Thomas, M. Munro, The end of the line for software visualisation?, IEEE Second Workshop on Visualizing Software for Analysis and Understanding (VISSOFT). IEEE Press, 2003.

[4] D. Cheriton, Man-machine interface design for timesharing systems, in Proceedings of the Annual Conference ACM76, pages 362–366. ACM Press, 1976.

[5] I. Chikofsky, J. H. Cross, Reverse Engineering and Design Recovery: A Taxonomy, *IEEE Software*, vol. 7, no. 1, 1990.

[6] S. Diehl, *Software Visualization - Visualizing the Structure, Behaviour, and Evolution of Software*. Springer, 2007.

[7] L. Fu, Exploration and Visualization of Large Execution Traces, Master of Computer Science Thesis, University of Ottawa, June 2005

[8] B. Gaines and P. Facey, Some experience in interactive system development and application, in Proceedings of IEEE no. 63(6), pages 894–911. IEEE Press, 1975.

[9] K. Gallagher, A: Hatch, and M. Munro, Software Architecture Visualization: An Evaluation Framework and Its Application, *IEEE* on Advanced Visual Interfaces (AVI 2000), 2000, pp. 21-28.

[10] T. Kennedy, The design of interactive procedures for man-machine communication, in *International Journal of Man- Machine Studies*, 6:309–334, 1974.

[11] A. Lund, Expert ratings of usability maxims, Ergonomics in Design, 5(3):15–20, July 1997.

[12] J.I. Maletic, A. Marcus, and M.L. Collard, A task oriented view of software visualization, *Procs. IEEE First International workshop on Visualizing Software for Understanding and Analysi*s. 2002. IEEE Computer Society Press. 32-40.

[13] E. Mealy, D. Carrington, P. Strooper and P. Wyeth, Improving Usability of Software Refactoring Tools, Proceedings of the Australian Software Engineering Conference (ASWEC 2007), pp. 307-318

[14] J. Nielsen (1994). Heuristic evaluation. In Nielsen, J., and Mack, R.L. (Eds.), Usability Inspection Methods, John Wiley & Sons, New York.

[15] Pacione M.J., Roper M., and Wood M.. A novel software visualisation model to support software comprehension, in Proceedings of the 11th Working Conference on Reverse Engineering, pp. 70–79. 2004. Delft.

[16] M. Petre, E. De Quincey, “A gentle overview of software visualization”, PPIG Newsletter, 2006.

[17] B.A. Price, I.S. Small, and R.M. Baecker. A Taxonomy of Software Visualization. In Proc. of the 25th Hawaii International Conference on System Sciences (HICSS), vol. 2, 1992, 597-606.

[18] B. Schneiderman. Designing the User Interface: Strategies for Effective Human-Computer Interaction. Addison-Wesley, 1987

[19] S. E. Sim, M. D. Storey (2000), A Structured Demonstration of Five Program Comprehension Tools. Lesson Learnt, In Proceedings of the Seventh Working Conference on Reverse Engineering (WCRE'00)

[20] J. Stasko, J. Domingue, M. Brown, B. Price. *Software Visualization*. MIT Press, 1998.

[21] M.-A.D. Storey. On Designing an Experiment to Evaluate a Reverse Engineering Tool. Proceedings of the 3rd Working Conference on Reverse Engineering (WCRE '96) Page: 31.

[22] M.-A.D. Storey, S. Cubranic, and D.M. German, On the use of visualization to support awareness of human activities in software development: a survey and a framework, in Procs. of the 2nd ACM Symp. on Software Visualization, pp. 193-202, May 2005.

[23] T. Systä, K. Väänänen-Vainio-Mattila, On empirical studies to analyze the usefulness and usability of reverse engineering tools, In Workshop Empirical Studies in Rev. Eng., STEP 2005, Budapest, Hungary.

[24] A. A. Terekhov, Automating Language Conversion: A Case Study, in Proceedings of 17th IEEE International Conference on Software Maintenance (ICSM'01), pp. 654-658.

[25] S. R. Tilley. Domain-retargetable Reverse Engineering II: Personalized user interfaces. In International Conference on Software Maintenance (ICSM '94)

[26] P.Tonella, M. Torchiano, B. Du Bois, T. Systä, Empirical studies in Reverse Engineering: state of the art and future trends, *Empirical Software Engineering* 12(5): 551-571 (2007).

[27] A. Von Mayrhauser, A. Marie Vans, Program Understanding Behavior During Debugging of Large Scale Software, Proceedings of 7th workshop on Empirical Studies of Programmers, 1997.

[28] A. Walenstein, Theory-based Analysis of Cognitive Support in Software Comprehension Tools, in Proceedings of the 10th International Workshop on Program Comprehension, 2002 ISBN:0-7695-1495-2

[29] Young, P., and Munro, M. (2003) Visualising software in virtual reality. IEEE First International Workshop on Visualizing Software for Understanding and Analysis. IEEE Computer Society Press.

[30] I. Zayour, T. C. Lethbridge, A Cognitive and User Centric Based Approach For Reverse Engineering Tool Design, Proceedings of the 2000 Conference of the Centre for Advanced Studies on Collaborative Research, Mississauga, Ontario, Canada, pp. 16, 2000.

A Methodological Framework to the Visual Design and Analysis of Real-Time Systems

Kawtar Benghazi, Miguel J. Hornos, Manuel Noguera, María J. Rodríguez
Universidad de Granada, ETSIIT Department of Software Engineering,
18071 Granada (Spain)
{benghazi, mhornos, mnoguera, mjfortiz}@ugr.es

Abstract

In this paper we present MEDISTAM-RT, a methodological framework for the design and analysis of real-time systems. We expose visual models and metamodels based on the UML-RT notation. These models are intended to capture different system views of a real–time system: contextual, static, dynamic and temporal views. We describe a systematic modeling process based on the aforementioned visual notations and illustrate its applicability by means of a real case study consisting of an environment control system for people with special needs.

1. Introduction

Visual modeling languages and techniques use graphical notations which are more natural, intuitive and understandable than textual notations. These languages and techniques have been increasingly adopted for the specification, design, development, and even validation of software systems [4] [11]. It is known that visual modeling technologies increase the abstraction level of design models, making them simpler and more understandable. Moreover, they facilitate the communication and collaboration among teams of designers, since models created with them allow graphically visualizing the structures and behaviors of the system (or subsystems) to be modeled. The Unified Modeling Language (UML)[6], which has been standardized by the Object Management Group (OMG), is the most used visual modeling language for software design and development. UML offers strong support for requirements engineering and system architecture, since it provides constructs for specifying the behavior and structure of a system through partitioning and hierarchical decomposition.

UML provides extension mechanisms to address its own limitations. One example is UML profiling, which adapts UML to a specific domain. Another example is metamodeling, which can be used to extend UML itself.

The UML-RT profile[9] is becoming a standard for designing visual models of complex real–time systems. UML-RT extends the *de-facto* standard UML with additional concepts (capsules, ports, protocols, and connectors) that provide an important support to model the static architecture of real-time systems. However, as for modeling the dynamic aspects of these systems it lacks of a syntactic model which allows communication between different interactive components of a real-time system to be coordinated. Moreover, although UML was designed for a real-time systems, it does not provide any support to deal with time issues (temporal restrictions, execution time). Additionaly, the semiformal semantic of UML-RT represents a considerable limitation, specially in real-time system design, in which the verification of properties, such as deadlock freeness and liveness, is essential. Moreover, the imprecise semantic of UM-RT models increases the risk of inconsistencies between models along a design process.

By observing the limitations which presents the application of the standard to the real-time system development, and focusing on the deficiencies of temporal constraint expressions, we consider that approaches that overcome these limitations should:

1. Extend the UML-RT syntax: with modeling concepts and annotations related to timed issues to improve the expressive power of UML-RT w.r.t. the modeling of dynamic aspects.
2. Specify the semantics of UML-RT: the temporal constraints must be supported by a formal semantics, which has a mathematical principals and foundations that allow system specification to be validated and verified.

In order to preserve the ease of use of UML-RT and strengthen its capacity in managing the complexity of real-time systems, we have extended it with a set of models that include temporal annotations inspired in the formal method CSP+T and defined a set of rules to transform models based on UML-RT into timed traces. All this is defined within MEDISTAM-RT, a design and analysis framework presented herein.

There exist several methodologies [10] [7] [5] to model real–time systems. However, our approach differs in that it adopts a Software Engineering oriented perspective combining semi-formal methods based on UML-RT with a formal language, in such a way it facilitates the design of complex systems and at the same time allows the verification and validation of the system specification.

This paper is structured as follows: Section 2 presents the multiviews of MEDISTAM-RT and the models based on UML-RT defined for each of them. Section 3 describes the modeling process and its application to a real case study and explains the way to formalize the MEDISTAM-RT models. Additionally it explains the use of timed sequence diagram as a visual tool to establish the consistency between system models. Conclusion and future work are summarized in Section 4.

2 MEDISTAM–RT Views

The methodological framework MEDISTAM–RT [2], provides a set of models based on UML-RT. The basic structural element used in MEDISTAM-RT is the *capsule*, which represents components at different level of abstraction. The capsule communicates with the exterior through *ports* and connect with other capsules by means of *connectors*. The signals interchanged between capsules are encapsulated in *protocols*. Moreover, the design of real-time systems in MEDISTAM-RT is modeled in base of three views, each of which will be described in the next subsections.

2.1 Static View

The static view describes the internal structure and properties of the system under two perspectives:

- *Structural*: it details the parts of each component and how they are related.
- *Data*: it describes the data handled in each component.

This view is modeled with two complementary diagrams: the *class diagram* and the *composite structure diagram*, which are described in detail below.

2.1.1 MEDISTAM–RT Class Diagram

The class diagram describes the external structural aspect of a system, and determines the elements that characterize the system, as well as the relationships between them. The elements or capsules of the system are classifiers which act as data containers, and communicate with other capsules through the interchange of events according to a protocol established between them (see example in Figure 1).

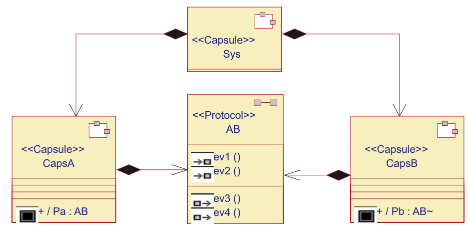


Figure 1. *Example of a MEDISTAM–RT class diagram*

These diagrams allow covering both structural and data aspects, providing essential importance to the communication interchange that occurs between the structural elements. In this view, the structural elements act like black boxes, hiding their data and behavior. Figure 2 shows the metamodel of MEDISTAM-RT class diagram.

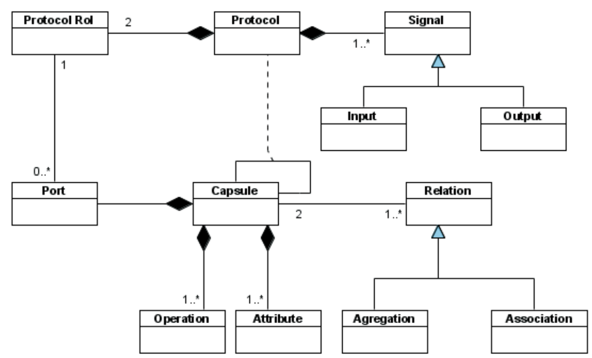


Figure 2. *Metamodel of MEDISTAM-RT class diagram*

2.1.2 Composite Structure Diagram

From a structural point of view, the composite structure diagram represents the internal organization of a capsule (or the system) as a composition of sub-capsules that are connected with each other to provide a service to the capsules that communicate with it. Unlike class diagrams, the capsule is seen here as a white box, and therefore the composite structure diagrams specify the internal consequences for any service request from outside by identifying the internal communication channels and sub-capsules that have to respond to that request (see an example in Figure 3).

The composite structure diagram of MEDISTAM-RT facilitates the architectural design of a system. Figure 4 shows

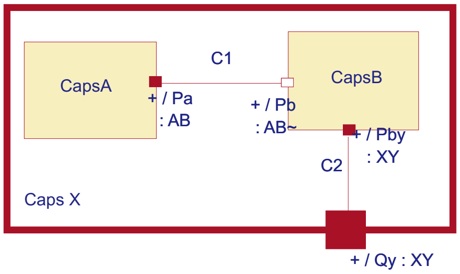


Figure 3. *Example of a composite structure diagram*

the metamodel displaying the modeling elements that each composite structure diagram may include.

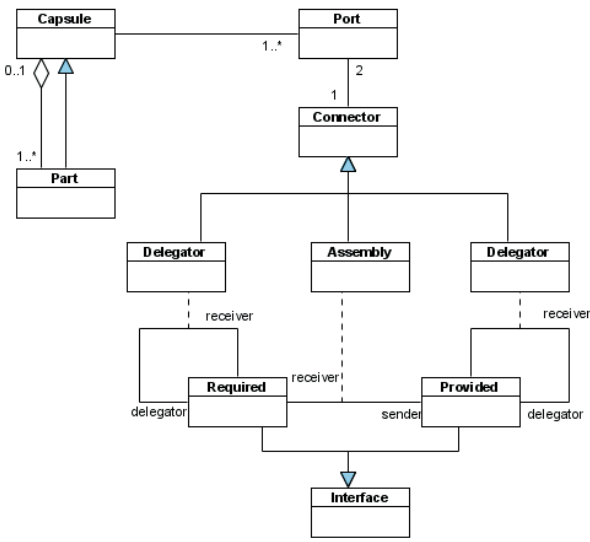


Figure 4. *Metamodel of the composite structure diagram*

2.2 Contextual View

We need a diagram that allows showing the list of messages representing the interaction between actors and the system. UML Use Case Diagrams lack of the expressive power to capture this view. To fill this gap, different pieces of work [3] have proposed the use of context diagram from a traditional structured method in UML based design.

In our approach, we use UML-RT class diagrams at different levels of abstraction as a kind of context diagram. The signals interchanged between the capsule and each actor, are represented by arrows with the same communication directions (see an example in Figure 5).

We believe this diagram combines the scope of both traditional context and class diagram, in such a way that the

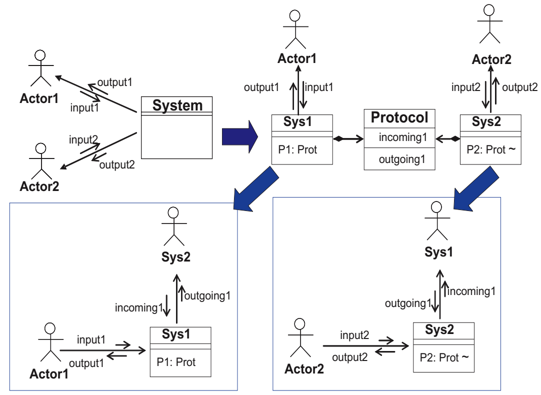


Figure 5. *Context diagram at different level of abstraction*

context diagram can be used to show a unified view of the system context and its internal view (protocols).

2.3 Dynamic and Temporal View

This is a dynamic and temporal view that allows the system to be visualized as a black box or a white one. These different perspectives make it possible to distinguish the external behavior of any capsule from the internal one.

- *Internal temporized behavior*: This aspect, which gives a white box view of the capsule, describes how the capsule behaves to provide services or carry out tasks.
- *External temporized behavior*: This aspect, which provides a black box view of the capsule, is concerned with the capsule interaction with its environment. It represents an abstraction of the services provided or required over time by the capsule via its ports (interfaces), by hiding all the internal events. In other words, this aspect represents an abstraction of the internal temporized behavior.

2.3.1 Timed State Diagram (TStD)

UML state diagrams are graphical tools that allows the description of the system behavior. To extend the scope of these diagrams so that they can properly specify the temporal behavior of real-time systems, we have enriched their syntax with new notions inspired in the syntax of CSP+T language [1]. We call this extension Timed State Diagram (TStD)[**?**]. Figure 6 depicts the metamodel for this type of diagrams.

[1]CSP+T is known for its flexibility to describe the dynamic aspect of systems, as well as for its temporal properties.

1. *Instantiation event*: it represents the origin of a global time at which processes can start their execution

2. *Marker event*: it is an event used jointly with a marker variable to record the instant at which the event occurs. For example, the annotation $Ev/t = gettime()$, means that the time at which the event Ev is observed during a process execution is in the marker variable t.

3. *Restricted event*: it is an event associated with a time interval which represents the period of time over which the event cab performed. This interval can be absolute (the event) or relative.

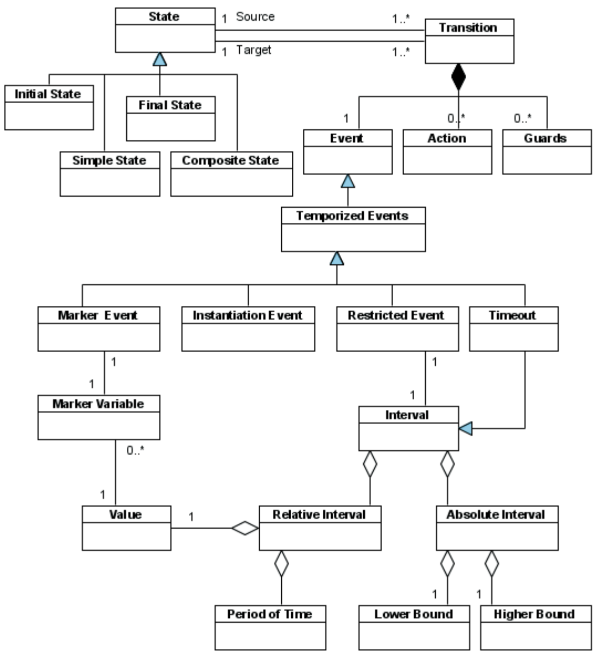


Figure 6. *Metamodel of TStDs*

2.3.2 Timed Sequence Diagram (TSeD)

Sequence diagrams show how the objects or other classifiers of the system interchange messages to carry out a task. These diagrams can be applied at different levels of detail, but they can be mainly used at the detailed design phase, where a communication between capsules has to be established according to predefined protocols.

On the one hand, the interaction of each capsule with its environment is carried out by through its interfaces (called ports), which represent the controller part of the capsule. This means that the ports are active elements that have their own behavior. Additionally, each component (i.e., capsule) has different independent ports that can operate in parallel. In other words, the communication of a component with its

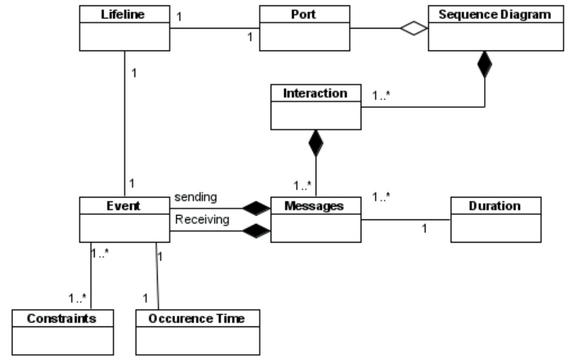


Figure 7. *Metamodel of TSeDs*

environment is rather oriented towards a vision of the interactions between each of its ports and its environment. For this reason, sequence diagrams in MEDISTAM-RT describe the interaction between capsules ports, which are considered active processes. This allows CSP handshaking principle for the interchange of messages to be used.

On the other hand, to establish a secure and correct communication, it is essential taking into account temporal aspects, which are assigned to the capsule behavior and which allow specifying time constraints. We have therefore enriched the UML sequence diagrams with temporal annotations, which allow specifying temporal constraints assigned to the interactions between capsules. We call Timed Sequence Diagram (TSeD) to this extension of UML sequence diagram. Figure 7 represents the metamodel that defines this extension. TSeDs represent the external behavior of components, i.e. its interactions with its environment. These interactions are based on messages. Each message is associated with:

1. A transmitting port and a receiving port.

2. A sending event, which represents the action of sending the message by the transmitter, and a receiving event, which represents the action of receiving the message by the receiver. Each event is associated to an occurrence time.

3. Duration. Time that goes by from sending to receiving the message.

4. Constrains, i.e. partial order between events and temporal constraints.

3 System Modeling Process

In this section, we will introduce the software engineering strategies used along the modeling process and the detailed steps to design real-time systems using the extended

UML-RT diagrams presented in the previous section. In addition, we will apply them to a real case study: the environment control for people with special needs.

It should be noticed that in the last years our research group have been working in the field of augmentative and alternative communication for people with cognitive disabilities (autism, dysphasia, cerebral palsy, etc.); in fact, we have developed the *Sc@ut* communicator [8]. We are trying to extend the functionalities of this portable communicator, in order to enable people with physical impairments to control diverse elements of a room (e.g., windows, doors, lamps, etc.).

In this context, with mentally and/or physically challenged people, dependable *firefighting systems* are critical. For a prompt action in case of fire, the action plan must be automatically controlled by this system and supervised by a tutor, since people in the room may not notice the fire (due to their cognitive disabilities) or have problems to escape from it (due to their physical disabilities). They need help to identify the risk (e.g., by means of a loud alarm) and escape (e.g., by means of presence detection and automatic door opening mechanisms). The plan should also include the activation of the fire sprinkler system, but before this, it is important to notify the alarm activation to the tutor and give him a certain time to cancel the action plan (in case of false alarm); this will mitigate individuals' stress.

Consequently, we will select the firefighting system and the door of a room to show in this section the models we have designed for these components by applying our methodology to this real case study.

3.1 Hierarchical decomposition

Due to the complexity of real-time systems, the system decomposition becomes a crucial step in the system design and analysis. Decomposition is a form of refinement, since each step in a hierarchical decomposition will allow identifying the internal structure of the composite component in question, and implicitly defines the connections among its subcomponents, which represent the system in more detail. The outcome of the decomposition implicitly reflect the solution proposed by the designer to resolve the problems, since in this step the designer makes decisions about which components must collaborate and also assumes the responsibility of setting some tasks to specific components (subcomponents).

3.2 Models Transformation: Abstraction and Refinement

Abstraction and refinement are two fundamental and complementary principles applied to systems design. Abstraction consists in hiding or removing irrelevant information to improve the understandability of the problem domain and to simplify the evaluation of the proposed design solution. However, refinement consists in adding more details to the system design, to be closer to the final solution, that is, the system implementation. Precisely, each abstraction step tends to answer more clearly the question "*what* must the target system do?" and each refinement step tends to answer to "*how* must the system behave to reach its objectives?"

3.3 Top–Down Modeling Process

The design process begins by decomposing the system into a set of subsystems, as depicted in Figure 8. Each subsystem may be modelled independently later on. Moreover, each subsystem may also be decomposed until reaching an adequate refinement level [2].

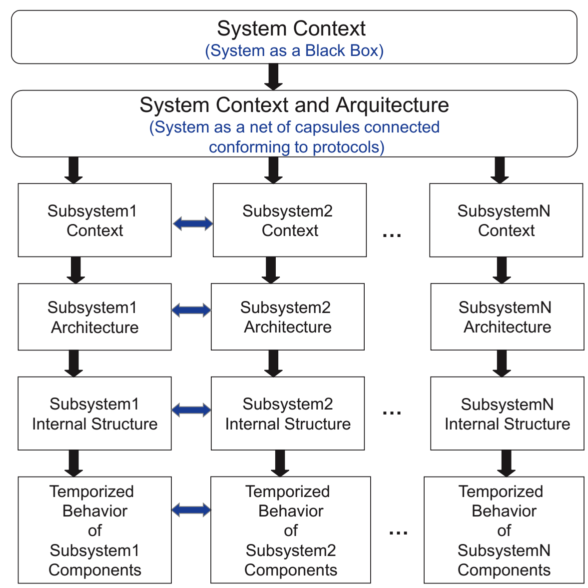


Figure 8. *Structure of the system modeling process of MEDISTAM–RT*

The system may be considered as a net of capsules that are connected by ports and that communicate each other according to predefined protocols. This net is modeled using a *context diagram*, like the one in Figure 9. This figure shows the complete system architecture of the selected case study and the interactions between the actors and subsystems. The system has been divided into two subsystems: *Door* and *Firefighting-system*. The functioning of the system is as follows: When smoke is detected in the room,

[2] This is the refinement level where all the subsystems are primitives.

the system activates the loud alarm immediately and sends a notice to both the alarm reception central and the tutor. At the same time, the presence detector checks if there are some people in the room; in this case, the door is automatically opened, allowing the occupants to leave. If the room is empty, the door must be closed, to try to avoid the spreading of the (possible) fire. Since the instant of smoke detection, the system waits two minutes before the sprinklers start to pour water (to prevent a false alarm, which must be communicated by the tutor to the system). If after that time, no signal of false alarm has been received by the system, sprinklers start to operate.

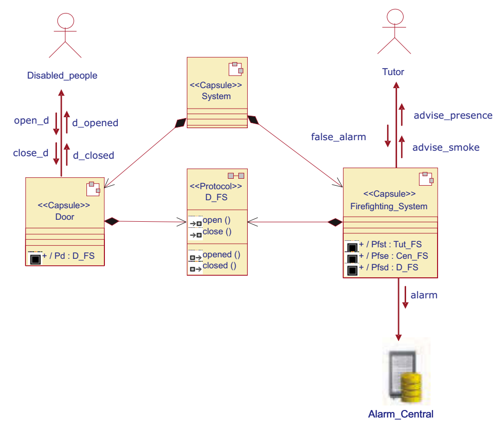


Figure 9. *Context diagram showing the complete system architecture*

3.3.1 Context Subsystem Modeling

When modeling a subsystem, we consider all its external elements as belonging to its environment. In this way, each subsystem interacting with the subsystem in question is considered by this as an external system.

The subsystem context is captured by a context diagram, where a subsystem is presented as a black box, and the subsystems interacting with the subsystem in question are modeled as actors. For example, in Figure 9, the subsystem *Door* interacts with the actor *Disabled-people* and the subsystem *Firefighting-system*. Likewise, the context diagram of the subsystem *Firefighting-system* interacts with the actors *Alarm-central* and *Tutor*, and with the subsystem *Door*, but in this case all of them as represented as actors, as Figure 10 shows.

Additionally, communication between subsystems and actors is encapsulated in communication protocols. And ports must also be created to communicate with the actors. For example, the port *Pst* and the protocol *Tut-FS* have been created in Figure 11 to encapsulate the communication between the *Firefighting-system* and the *Tutor*.

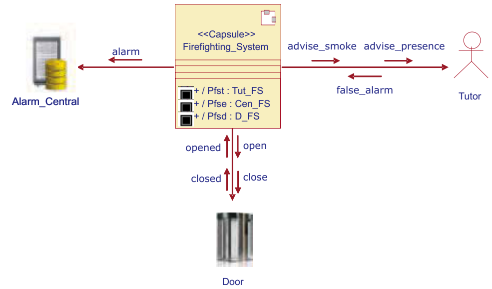


Figure 10. *Context diagram of Firefighting-system*

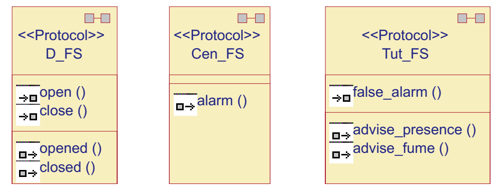


Figure 11. *Communication protocols between Firefighting-system and its environment*

3.3.2 Subsystem Architecture Modeling

After dividing a subsystem into lower level subsystems, we model its architecture by means of a class diagram, which shows the components of the subsystems and the signals they interchange. For example, Figure 12 shows the architecture of *Firefighting-system*.

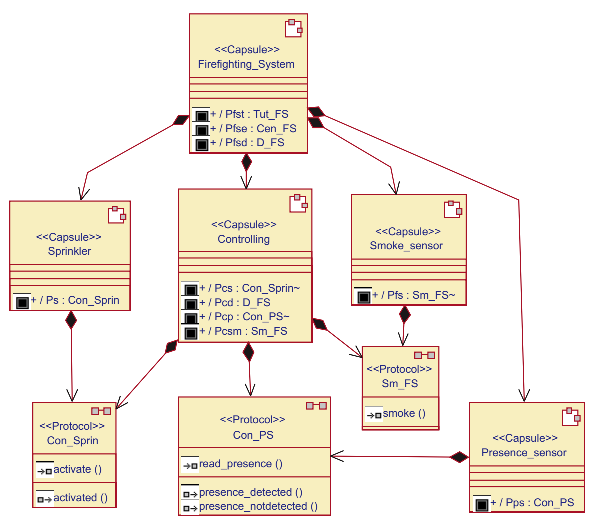


Figure 12. *Class diagram of Firefighting-system*

3.3.3 Internal Structure Modeling

The internal structure of each subsystem is described by a composite structure diagram, which shows the subsystem components and how they are connected. The internal structure of *Firefighting-system* is shown in Figure 13.

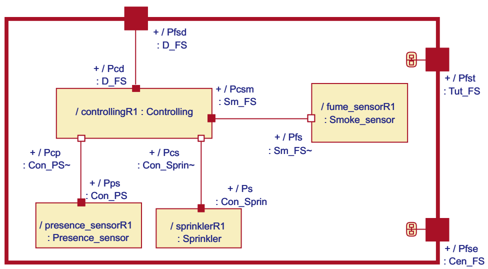


Figure 13. *Internal structure of Firefighting-system*

3.3.4 Subsystems Temporized Behaviour Modeling

The last modeling step consists on defining the temporized behavior of subsystems [3] using TStDs. Figure 14 shows the timed behavior of the component *Controlling* of the subsystem *Firefighting-system*.

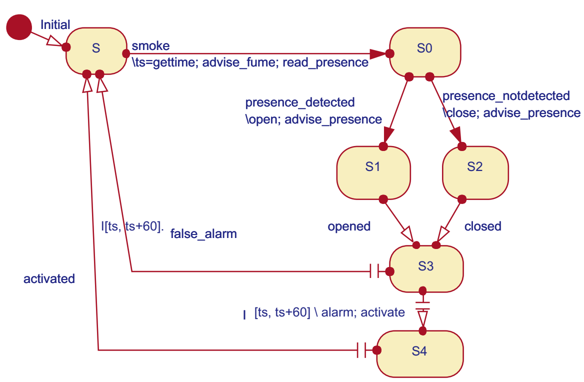


Figure 14. *TStD of the component Controlling*

3.3.5 Formalization and Consistency Verification of MEDISTAM–RT Models

Rigorous specification and analysis of system properties requires the integration of formal methods into the system development process. In our methodology, we use a transformation-by-derivation technique that consists on defining a set of transformation rules that maps systematically MEDISTAM-RT models into formal models. Likewise, a set of rules has been defined to transform class diagrams, composite structure diagrams and TStDs into CSP+T processes, and other rules have been defined to transform TSeDs into timed traces. Figure 15 shows a scheme of these transformations.

[3] The primitives subsystems are directly represented by TStDs, since they have no internal structure to describe.

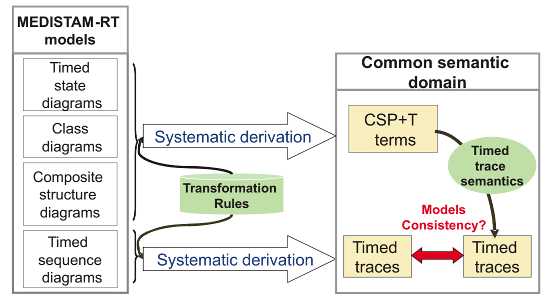


Figure 15. *Models Transformation in MEDISTAM–RT*

Additionally, to open up the possibility of verifying the consistency between the different views represented by the different diagram we map all diagram to a common semantic domain, timed traces.

3.4 TSeD as Visual Tool to Establish System Models Consistency

The multiviews of MEDISTAM-RT are closely intertwined, since they have a set of related and common elements. The protocols in a class diagram encapsulate the events which pass through the ports and connectors defined in the corresponding composite structure diagram. Moreover, the behavior of the components, which is described by TStDs, must satisfy the restrictions required by the protocols assigned to their ports (i.e., temporal constraints and the message order), as Figure 16 shows. In addition, a TSeD, which is a structured representation of the protocol behavior as a series of sequential steps over time, shows possible interactions between capsules. Actually, the functionality of a safe (i.e., deadlock free) system can be assured if the temporal consistency is established among the TStDs corresponding to the ports of communicated capsules. This is a very hard task if the design of the TStDs has to be done by reasoning separately on each capsule.

Furthermore, a TSeD representing the behavior of a given protocol gives a general view of the interaction between capsules, and determines the pattern of interchange of messages between the involved ports, which belong to different capsules. Hence, the TSeD can be used as a tool to schedule the message passing as well as to specify the time restrictions on the message interchange between a capsule

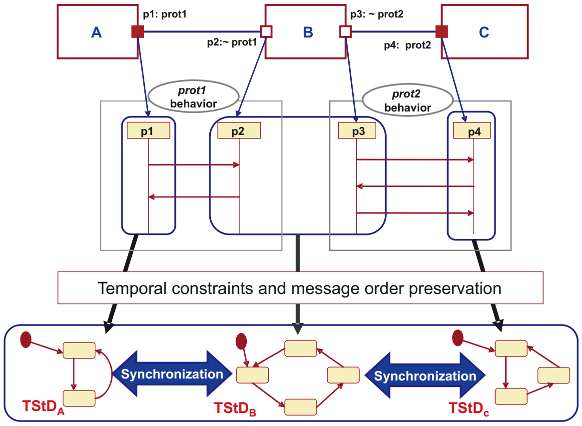


Figure 16. *Using TSeDs for establishing components consistency*

port and its environment. In this way, a capsule is correct if the temporal constraints and message order in the TSeD of the protocol is maintained in the TStD of the corresponding capsules.

4 Conclusions and Future Work

We have presented a series of tools intended to the design and analysis of complex systems and the communication and collaboration among project team members. These tools are part of our methodological framework MEDISTAM-RT, which extends UML and UML-RT diagrams in order to adapt them to the modelling of real-time systems. We have also presented a systematic modeling process based on these diagrammatic notation and illustrate its applicability by means of a real case study consisting of an environment control system for people with special needs.

As future work, we aim at extending our framework with a new views, i.e. Requirement and implementation and complete our process modelling in the way that it covers the full lifecycle of software development.

Acknowledgments.

This research is funded by the Spanish Government's Ministry of Science and Innovation, via the project TIN2008-05995/TSI.

References

[1] Benghazi, K., Capel, M.I., Holgado, J.A., Mendoza, L.E.: A methodological approach to the formal specification of real-time systems by transformation of UML-RT design models, Science of Computer Programming 65(1), 41-56 (2007)

[2] Benghazi, K.: MEDISTAM–RT: Metodología de diseño y análisis de sistemas de tiempo real. PhD Thesis, University of Granada, Spain (2009)

[3] Bourdeau, Philippe Lugagne, P., Roques, P.: Hierarchical Context Diagram with UML : An experience Report On satellite Ground System Analysis, The Unified Modeling Language. UML98: Beyond the Notation, (1999)

[4] Jiang, M., Groble, M., Neczwid, A., Willey, A.: Modeling real-time communication systems: Practices and experiences in Motorola, Journal of Visual Languages and Computing & Computing 17 (2006) 584–605

[5] Medina, j.: Metodologa y herramientas UML para el modelado y anlisis de sistemas de tiempo real orientados a objetos. PhD Thesis, Cantabra University, Spain(2005)

[6] OMG, UML Superstructure Specification – version 2.0, Object Management Group, Massachusetts, USA, (2004)

[7] Ramesh, S., Ganesan, P.V.V., Raravi, G.: A Formal Framework for the Correct-by-construction and Verification of Distributed Time Triggered Systems, In: International Symposium on Industrial Embedded Systems, pp. 63-70 (2007)

[8] Rodríguez-Fórtiz, M.J., González, J.L., Fernández, A., Entrena, M., Hornos, M.J., Pérez, A., Carrillo, A., Barragán, L.: Sc@ut: Developing Adapted Communicators for Special Education, Procedia - Social and Behavioral Sciences, 1(1), (2009)

[9] Selic, B., Rumbaugh, J.: Using UML for Modeling Complex Real-Time Systems, ObjecTime Limited/Rational Software whitepaper (1998)

[10] Sifakis, J.: A Framework for Component-based Construction Extended Abstract. In: Third IEEE International Conference on Software Engineering and Formal Methods, pp. 293-300, IEEE Computer Society, Washington DC, USA (2005)

[11] Kremer, R.: Visual Languages for Knowledge Representation. In Proc. of 11th Workshop on Knowledge Acquisition, Modeling and Management (KAW'98), Morgan Kaufmann (1998)

Visualizing Pointer-related data flow Interactions

Marcel Karam, Marwa El-Ghali, Hiba Halabi

American University of Beirut
{mk62 ,mma73,hah25} @ aub.edu.lb

Abstract

Visualizing definition and use interactions in a procedure; especially those related to pointers, is one way to help ease the tasks of inspecting the results of passing and failing test cases and subsequently locating pointer related faults. To assist with the fault localization of pointer-related faults, we present in this work a visualization framework that makes it possible to account for the types of definition and use interactions that are relative to pointer references, and generate insightful visual artifacts to show the various types of pointer interactions.

1. Introduction

It's known that testing and locating errors are some of the most time-consuming tasks of software development. One way of representing a program for the purpose of testing is to transform its code into a graph where sequences of statements are represented as nodes and decisions are represented as edges. To test a program based on this graph, several test adequacy criteria can be applied. For example, *all-nodes* and *all-edges* try to cover every node and every edge in the graph, respectively. More rigorous criteria such as *all-uses* tries to cover some unique paths in the graph containing definition and use associations (du-association or du-pairs) from a node containing a definition of a variable to all nodes containing its uses. The latter criterion is particularly important as it helps to cover more paths combinations in the graph then *all-nodes* and *all-edges*, and subsequently additional insights about the data dependencies of a program under test. Several visualization techniques, for the purpose of fault localization, have been used to represent pass/fail test cases involving statements [3], nodes and edges [2], and non-pointer related dataflow interactions [5]. With languages that support pointers however, visualizing dataflow interactions becomes increasingly complex as variables can be either used or defined without an explicit code statement, i.e. through pointer dereferences, and pointers aliases sets or the set of other pointer variables that point to the same memory location. The lack of efficient visualization techniques that handle the visualization of pointers interactions and their types [9, 11] (i.e definite definition or possible use) for the purpose of testing, makes the task of obtaining textual information from a testing criterion such as *all-uses* very difficult to interpret by testers. To help end-users or testers with this task, we have developed a sophisticated visualization framework that accounts for intra-procedural du-pairs types (see Table 1) that are relative to pointer references, and used insightful visual artifacts to show intricate pointer interaction details. The remainder of this paper is organized as follows. Section 2 presents related work. Section 3 gives an overview of the techniques we have used to collect the pointer variable interactions and their alias sets. Section 4 explains the type–based pointer visual artifacts used in our visual testing framework; especially complex pointers interactions. Section 5 concludes with future work.

2. Related Work

In the direction of graph or visual artifacts representation of a program under test for the purpose of fault localization, few techniques have been proposed. The work in [2] implemented a technique to show the control flow graph (using circles and arrows), and as such limiting the screen in the amount of information it can show. The work in [3] implemented tools where each statement in a program under test is represented as a pixel. The pixels are then colored using a continuous coloring scheme that colors these pixels (from complete red to complete green) according to their participation in passed/ failed test cases. This approach maximizes the use of the screen in showing the tested statement as pixels but limited the visualization of more complex information such as dataflow interactions. The work in [5] implemented a framework to visualize for a function under test, the blocks, loops, edges, and non-pointer du-associations. A non continuous coloring scheme was used to visualize the results of a testing session. The framework allowed a tester to perform *all-nodes*, *all-edges*, and *all-uses* testing; thus expanding fault localization to include not only source code statements or edges, but also data-flow interactions. In this work we extend the framework in [5] to support pointer-type dataflow interactions; making it a first of its kind to visualize du-associations related to pointer and their alias sets, as well as represents their participation of in pass/fail test cases.

3. Pointer Analysis

Pointer analysis precision is of paramount importance in the field of dataflow testing. In our visualization framework we employ the pointer analysis algorithm in [1] to collect all the aliases for each pointer in a function under test. We then account for the types of du-pairs and use them to assist the tester in both coverage and fault-localization; a practice that has been neglected by many researchers. Thus, we compute data dependences and classify them based on types of definitions, uses, and paths from the definitions to the uses. According to the work in [9], definitions and uses are considered as either definite or possible, based on whether they can be determined statically to refer to a single variable or a set of variables. Then, the set of paths related to each pair of definitions and uses is classified based on its def-clear status, i.e. whether it contains redefinitions or not. This classification is an indication of the strength of the binding of the du-association that is derived from these definitions and uses. Sets of reaching-definitions are computed by using data-flow equations, and then reaching definitions for each use are examined [9]. Next, definition-use (def-use) associations, which relate statements that assign values to variables, to the statements that use these variables, are computed by using the reaching definitions and uses equations [4].

Table 1. The du-pairs and their types

I	definite def (ddef) definite use (duse)
II	definite def (ddef) possible use (posuse)
III	possible def (posdef) definite use (duse)
IV	possible def (posdef) possible use (posuse)

Based on the types of the definition, use, the paths occurring between them, the def-use associations are classified into 24 categories [9]. However, in our work, we group the 24 types into four types for the purpose of efficient visualization. Table 1 lists the du-pairs and their types.

4. Visualization Framework

Our visualization framework takes as input, the source code of a C program and the pass/fail results of executing the source code with each test case in the test suite in the form of an excel© sheet. The framework then allows the users to perform *all-uses* at the intra-procedural level and track the dataflow interactions (pointer variables and their aliasing sets) that participated in the *pass* and *fail* tests of a test suite. We next present the specifics of our visual testing framework.

4.1. Graphical Representation

Our visualization framework offers the tester, in addition to viewing blocks, controls, and basic dataflow interactions, the capability to view intraprocedural and interprocedural du-pairs that are related to pointer variables and their aliases sets. We next describe the basic visual artifacts that represent blocks, controls and variable definition and uses, and then explain the visual representation of complex variable (pointers and dereferencing) interactions. This is as a major contribution as our framework previous work only allowed basic pointer analysis and visual representation. For a function in a program under test, a non-predicate basic block or a set of sequential statements is represented as a visual node 8-9 with a label representing the line numbers in the textual code. For example, in Figure 1, the visual node with label “8-9” represents a basic block that corresponds to lines 8 and 9 of function *main*. A predicate block or a block containing a predicate statement is represented as 6-7 where the ✓ and ✕ correspond to the *true* and *false* outcome, respectively. For example, in Figure 1, the visual node that is labeled “6-7” represents the predicate block containing the statements in lines 6 and 7 in function *main*. A non-pointer variable definition reaching the end of a block is represented by a *root* or *source* on the bottom of the visual node and a use of that variable is represented by a *terminal* or *a sink* on the top of the visual node. A du-pair is then represented as a datalink between the root and the terminal . For example the datalink in function *main* that is connecting the root representing *x* on the visual node that is labeled “1-4” to its terminal on the visual node that is labeled “5-6”, represents the du-pair with respect to variable *x* that is defined in line 1 and used in line 7 of *main()*. This datalink or du-pair is of type 1 in our visualization framework. Table 2 lists the five categories and their visual representations. In category 2, a pointer variable definition reaching the end of a block is represented with a (star-like shape root), and a pointer variable use is represented also with a , and a datalink connecting them correspond to a du-pair between two pointer variables. For example, in function *main*, the datalink between the definition of the pointer variable *p2* in the block that is labeled “1-4” and its use in the block that is labeled “8-9” represent the du-pair with respect to pointer variable *p2* that is defined on line 4 and used on line 8 through the statement “p3 = p2”. In category 3, a variable definition reaching the end of a block is represented with a root **,** a pointer dereferencing use is represented with a filled circle , and a datalink between them represents a du-pair between the definition of a variable and the use of it address its dereferencing non pointer variable. For example, in function *main*, the datalink between variable *y* in the block that is labeled “1-4” and its address use in the block that is labeled “8-9” represent the du-pair with respect to the definition of *y* on

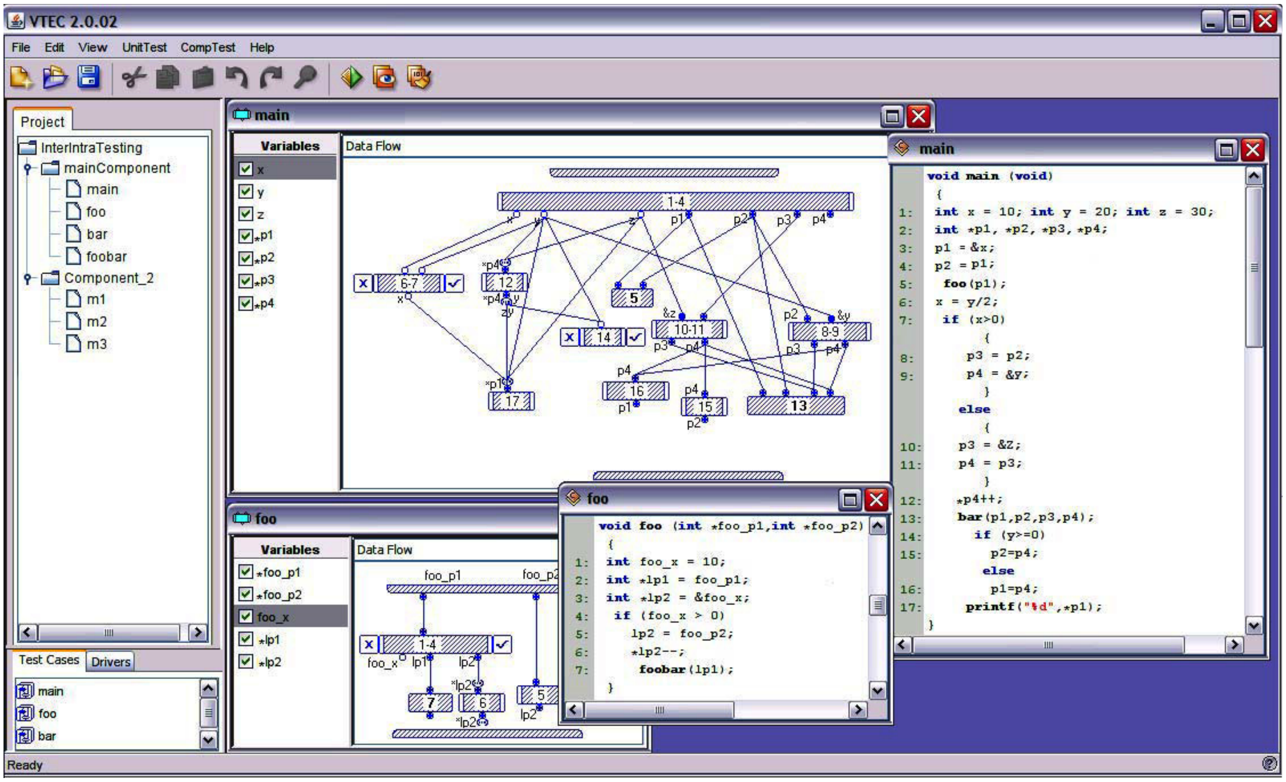


Figure 1. The Visual environment showing the dataflow interaction of pointers.

line 1 and the use of its address on line 9 through the statement "*p4 = &Y*".

Table 2. The 5 datalinks type (roots on left and terminals on right). Note: "C" in the table's header stands for Category.

C	visual	description
1		A datalink between the definition of non pointer variable and its use.
2		A datalink connecting a pointer variable definition to a pointer variable use.
3		A datalink between the definition of a non pointer variable definition and to its address use.
4		A datalink connecting two dereferenced pointer variables and their alias sets (from definition to uses according the type in Table 4.)
5		A datalink connecting a non pointer variable definition to a dereferenced pointer variable use. The list on this terminal type indicates the list of aliases.

In category 4, the definition(s) of a variable dereferencing an (alias set) is represented with a root, their use(s) at a variable dereferencing variable is represented with a terminal, and a set of datalinks representing du-pairs between them. For example, in function *main*, the set of aliases for **p4* (root on the visual node that is labeled 12) are possible definitions or *pdef* of *y* and *z*, and the set of aliases for **p1* (terminal on the visual node that is labeled 17) are possible uses *posuse* of *y* and *z*. Table 3 lists for each pointer in function *main* all its alias sets, and for each dereferencing variable the definitions and uses alias sets.

In category 5, a root representing the definition of a variable through, its use is represented with a, and a datalink between them represents a du-pair between the definition of a variable, and one of its possible aliases set uses. For example, in function *main*, the datalink between the roots representing the definitions of *x*, *y* and *z* on the visual node that is labeled "1-4" and their uses in the block that is labeled "17" represent the du-pairs with respect to *x*, *y* and *z* on line 1 and their uses as part of the alias set of *p1* on line 17 through the statement "printf (%d, *p1)". Table 3 lists the du-pair types we use in our visualization model. It also shows their types (Table 1).

4.1. Fault Localization and Color

Our framework uses the continuous coloring scheme that is used in [3]. It allows the tester to choose pointers or non pointer variables (through the checkboxes in Figure 1) or any variable and display its datalinks. This gives the tester the chance to focus on datalinks that have failed in using a testing criterion such as all-uses. When the tester chooses a pointer variable that has 2 aliases use at the same block, the framework makes the link blinks to draw the user's attention. For example, the datalinks between the blocks that are labeled "12" and "17" represent the following information <12, 17, y, IV>, <12, 17, z, III> (See Table 3). This means that one datalink represents 2 possible interactions. In our visualization framework this datalink blinks, and when the user puts the cursor on the link, a tooltip shows variables *x* and *y*.

Table 3. The format: <def line #, use line#, var, type>

main()	foo()
<1,6, y,I>	<1,4, foo_x,I>
<1,12, y,II>	<1,6, foo_x,II>
<1,14, y,I>	<0,6, *foo_p2,I>
<1,17, y,II>	
<6,17, x,II>	
<12,17, y,IV>	
<1,12, z,II>	
<12,14, y,III>	
<1,17, x,II>	
<12,17, z,III>	

5. Conclusion and Future Work

In this short paper, we presented a methodology which handles pointer variables by computing their corresponding alias sets, minimizing them by determining their equivalence classes, and classifying the occurring data dependences according to their strength types. We also extended the work in [5] and implemented a visualization technique that handles intra-procedural interactions in the presence of pointers and their aliases. The framework needs to use the coloring scheme that is used in [3] to representing uncovered program elements (pointer-related dataflow interaction) as well as the results of passing and failing tests. This coloring scheme, as in [3], could help in locating and uncovering faults. Also as future work, we could extend the framework to support inter-procedural testing of pointers interactions. We could represent static slices (for the purpose of testing) as visual slices which allow the user to further explore both intra- and inter-procedural data interactions. We could also further explore the slicing module to expand and refine it, possibly providing a separate visualization scheme in terms of views and coloring, especially on the interprocedural level. Also, the coverage component of our framework could be augmented to display the percentages of covered elements as well as those uncovered, relative to the total collection of elements, for the entire system as well as for each particular procedure in it. This could help the tester estimate, at a certain point in the testing procedure, the number of tests, time and effort still required to ensure testing coverage.

6. References

[1] Choi, J. D., Burke, M., Carini, P. "Efficient Flow-Sensitive Interprocedural Computation of Pointer-Induced Aliases and Side Effects". *In Procedings of the 20th ACM Symposium on Principles of Programming Languages*, pages 232-245, 1993.

[2] Harrold, M.J and Kolte, P. "Combat: A compiler based data flow testing system". *In Proceedings of the Pacific Northwest Quality Conference*, pages 311-323, October 1992.

[3] Harrold, M.J., Jones, J. A., Stasko. J. "Visualization of Test Information to Assist Fault Localization". *International Conference on Software Engineering Proceedings of the 24th International Conference on Software Engineering*, pages 467 – 477. Florida 2002.

[4] Hind, M., and Pioli. A. "Which pointer analysis should I use?" *In Proceedings of the. ACM Symposium on Software Testing and Analysis*, pages 113-123, 2000.

[5] Karam, M. R., and Abi Abdallah A. "Assisting in Fault Localization using Visual Programming Constructs". *In Proceedings of the 2005 Canadian Conference Electrical and Computer Engineering*, pages 856-860, May 2005.

[6] Landi W., and Ryder B. G. "A safe approximate algorithm for interprocedural pointer aliasing". In *Proceedings of the ACM SIGPLAN '92 Conference on Programming Language Design and Implementation*, pages 235–248, July 1992.

[7] Liang, D. and Harrold M.J. "Efficient points-to analysis for whole-program analysis". In *Proceedings of ESEC/FSE '99 7th European Software Engineering Conference and 7th ACMSIGSOFT Symposium on the Foundations of Software Engineering*, volume 1687 of *Lecture Notes in Computer Science*, pages 199–215. Springer-Verlag, Sept. 1999.

[8] Liang, D. and Harrold M.J. "Equivalence analysis: A general technique to improve the efficiency of data-flow analyses in the presence of pointers. In *Proceedings of the Workshop on Program Analysis for Software Tools and Engineering*, pages 39-46, 1999.

[9] Orso, A., Sinha, S., Harrold, M.J. "Effects of pointers on data dependences". In *Proceedings of 9th International Workshop on Program Comprehension*, pages 39-49, Toronto, Canada May 2001.

[10] Ruthruff, J., Creswick, E., Burnett, M., Cook, C., Prabhakararao, S., Fisher, M., and Main, M. "End-user software visualizations for fault localization." In Proceedings of the 2003 ACM Symposium on Software Visualization (San Diego, California, June 11 - 13, 2003). SoftVis '03. ACM, New York, NY, 123-132

[11] T.J. Ostrand and E.J. Weyuker. "Data flow-based test adequacy analysis for languages with pointers". In *Proceedings of the Symposium on Testing, Analysis, and Verification*, pages 74–86, Oct. 1991.

A Graphical Tool to Support Visual Information Extraction

Giuseppe Della Penna, Daniele Magazzeni and Sergio Orefice
Department of Computer Science
University of L'Aquila
Italy
Email: {giuseppe.dellapenna,daniele.magazzeni,sergio.orefice}@univaq.it

***Abstract*—In this paper we present the Spatial Relation Query tool, a graphical software system for Information Extraction based on the visual appearance of the information. The SRQ tool is provided with a query language similar to the well-known SQL, which allows users to write queries based on the visual arrangement of the information in an intuitive way. Two applications showing the use of the tool on web pages and geospatial data are presented.**

I. Introduction

The process of *knowledge acquisition* from generic information domains has its central phase in the *Information Extraction* (IE), which aims to extract from the located documents relevant information that appear in certain semantic or syntactic relationships. In particular, IE tries to process the relevant information found on the documents in order to make it available to structured queries. Most often, information extraction systems are customized for specific application domains, and require manual or semi-automatic training sessions.

IE from structured and semi-structured documents is frequently performed using *wrappers* whose most natural and widespread application area is the World Wide Web. Indeed, in the last years much research has been done for supporting information extraction on the web, and several approaches have been proposed. In particular, recent surveys of web data extraction tools [1], [2], [3] have singled out the following categories: *wrapper induction tools* (e.g., DEPTA [4]), *HTML-aware tools* (e.g., LIXTO [5]), *modeling-based tools* (e.g., DEByE [6]), *NLP-based tools* (e.g., WHISK [7]), and *ontology-based tools* (e.g., WeDaX [8]). Moreover, there are *specific-targeted tools* (e.g., WiNTs [9] or VENTex [10]) and *languages for wrappers development* (e.g., Minerva [11] or TSIMMIS [12]).

More similar to our approach are the works making information extraction based on the actual *visual appearance* of the information. For example, there are approaches that focus on specific application areas like, e.g., the work in [13], where the authors propose a machine learning methodology which allows to automatically extract specific field of PDF documents. Moreover, there are approaches based on *visual web page analysis* [9], [14], [10], [15], which exploit the visual web page representation. However, these works focus on information extraction specifically targeted on tasks like record boundary detection [9], web page segmentation [14], visual web table extraction [10], or visual similarities detection (e.g, the recognition of repetitive patterns) [15].

The major shortcoming of all the approaches above exploiting the visual appearance of information is the lack of an automatic counterpart supporting the visual information extraction.

In this paper we propose a general information extraction approach based on the visual appearance of the information, conceived as its user-perceived *rendering*. This allows to shift the IE problem from the low level of code (e.g., raster graphics, vector drawing, wordprocessor formatted text, web page, etc.) to the higher level of visual features, providing a paradigm of the kind "what you see drives your search" that supports a natural query formulation.

In particular, in our approach the extraction of information is driven by the spatial arrangement of the information (e.g., *spatial relations* such as "right of" or "included in"). The syntax of these spatial relations is based on the formal framework of *visual language classes* defined in [16]. Here, each class characterizes a family of visual languages based upon the nature of their graphical objects and composition rules.

To this aim, we have designed a SQL-like query language, namely *SRQL* (Spatial Relation Query Language), which allows to write queries based on the visual arrangement of the information in an intuitive way. Moreover, the SRQL language can also use further semantic and graphical attributes deducible from the information source to refine the queries, thus providing a framework where the information extraction can be performed integrating spatial relations, visual attributes, textual content and document structure. To give an idea, it will be possible to extract the tables on the right of an image, or the text below the first/last image, or the links between two text blocks within a web page, or the text of a certain color within a paragraph beginning with a determined word within a pdf file, etc.

The spatial relation formalism and the SRQL language have been implemented within the *SRQ* (Spatial Relation Query) software tool. The tool is provided with a graphical front-end which allows to define and manage a library of spatial relations, compose and execute queries, and export the extracted information for further analysis and manipulation.

The paper is organized as follows. In Section II we describe overall organization of the SRQ tool together with basic notions on the underlying spatial relation formalism and SRQL language. Section III contains two sample case studies, showing how the tool can be used to extract information from web pages and maps, respectively. Finally, some concluding remarks are outlined in Section IV.

II. The Spatial Relation Query Environment

Before describing the SRQ tool, let us give a brief overview on the formalism of spatial relations used in the system. More details on the theoretical aspects of this formalism can be found in [16].

A. Spatial Relation Theory

In general, the specification of the spatial composition rules is an important step in the design of visual systems, since a formal characterization of its structural characteristics is crucial to provide a systematic base and avoid ad-hoc implementations.

Basically, there are two modalities used for composition, i.e. by connecting or spatially arranging graphical objects. For the purposes of this work, we restrict our attention to the latter case and use a formal framework based on the *box spatial relations* as defined in [16]. In this framework the graphical objects are characterized by their bounding boxes, i.e., they can be syntactically described and manipulated through their bounding box, whatever the shape of the object is. Indeed, the spatial compositions that relate the graphical objects are defined on the coordinates of the upper-left and lower-right points of their bounding box.

The following definition gives a formal specification for the generic spatial relation REL with respect to a graphical object a. It uses two functions $UL_{REL}(m, n, m_1, n_1)$ and $LR_{REL}(m, n, m_1, n_1)$ that map the coordinates of the upper-left (m, n) and lower-right (m_1, n_1) points of the bounding box of a onto sets of points.

Given a graphical object a with coordinates of the upper-left and lower-right points (x, y) and (x_1, y_1), respectively, and a graphical object b with coordinates of the upper-left and lower-right points (h, k) and (h_1, k_1), respectively, then: $aRELb$ holds if and only if $(h, k) \in UL_{REL}(x, y, x_1, y_1)$ and $(h_1, k_1) \in LR_{REL}(x, y, x_1, y_1)$.

In other words, $aRELb$ holds if and only if the upper left and the lower right points of b are contained respectively within two areas calculated on the syntactic attributes of a through the functions UL_{REL} and LR_{REL}.

The above definition of composition yields three types of possible spatial arrangements: inclusion, intersection and spatial concatenation, where the term "concatenation" refers to any spatial arrangement of graphical objects not intersecting their areas.

As an example, the relations INCLUDE, UP, DOWN, LEFT and RIGHT, taken from [16], are defined in Figure 1. These relations model the general types of *inclusion* and *spatial concatenation*.

INCLUDE	$UL_{INCLUDE}(x, y, x_1, y_1) = \{(m, n) \mid x \leq m < x_1, y_1 < n \leq y\}$ $LR_{INCLUDE}(x, y, x_1, y_1) = \{(m, n) \mid x < m \leq x_1, y_1 \leq n < y\}$
UP	$LR_{UP}(x, y, x_1, y_1) = \{(m, n) \mid n > y\}$
DOWN	$UL_{DOWN}(x, y, x_1, y_1) = \{(m, n) \mid n < y_1\}$
LEFT	$LR_{LEFT}(x, y, x_1, y_1) = \{(m, n) \mid m < x\}$
RIGHT	$UL_{RIGHT}(x, y, x_1, y_1) = \{(m, n) \mid m > x_1\}$

Fig. 1. Definition of spatial relations

B. The SRQ Tool

The overall organization of the SRQ tool is shown in Figure 2. The current prototype application has been developed in Java with the support of *ANTLR* [17] for the spatial relation compiler and query interpreter. The application uses a plugin system to interface with several different *visual data sources* and build the corresponding box model.

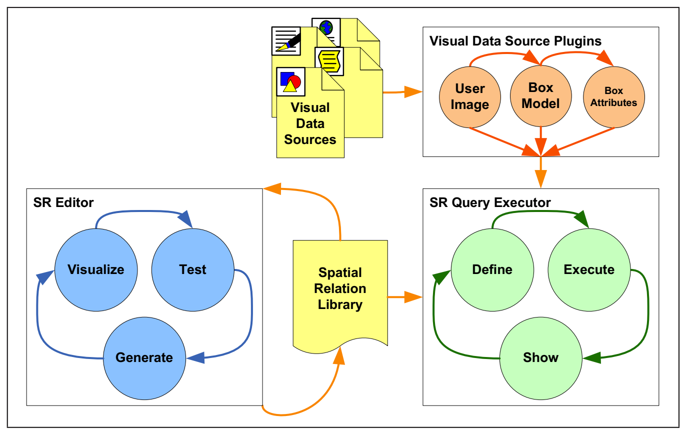


Fig. 2. SRQ tool architecture

In particular, the SRQ tool is composed of two interacting modules: the *Spatial Relation Editor* and the *Spatial Relation Query Executor*. These modules share the access to a *spatial relation library* which contains the basic building blocks of the query language.

The Spatial Relation Editor module allows to create and manage the spatial relation library, which defines the spatial relations used in the queries. The library itself is written in an appropriate *Spatial Relation Definition Language* (SRDL) that, for sake of simplicity, will not be described here. In particular, the user can exploit the following functionalities to manage the library through this interface.

Visualize This functionality allows to interactively visualize and explore the *meaning* of a particular spatial relation by graphically analyzing it.

Test This functionality allows to draw some graphical objects and test which relations are satisfied by their spatial arrangement.

Generate This functionality allows to generate new spatial relations through an assisted create-by-example process.

These functionalities will be also exploited within the Spatial Relation Query Executor, in order to support the composition of the queries. In particular, they will help the user to choose the appropriate relations (through the Visualize and Test functionalities) or create new ones (through the Generate functionality), if needed, during query session.

The Spatial Relation Query Executor allows to compose and execute queries. The composition of queries can be accomplished through the *Spatial Relation Query Language* (SRQL) language, which exploits the spatial relation theory given above to create queries based on the visual arrangement of the data and is designed to be as similar as possible to the well known SQL.

Before applying an SRQL query, the tool analyzes the information resource and generates the corresponding *user image*, i.e., an image which shows how the information actually looks like to the user. Then, it builds an appropriate *box model* that encapsulates all the distinguished objects from the visual appearance, possibly associating a set of *attributes*, deducible from the source information, to each box. Depending on the particular IE data domain, such attributes characterize specific properties of the objects, which may be both visual, (e.g., color of the box, text contained in the box, etc.), and semantic (e.g., if the box has a particular predefined meaning in the information).

The basic structure of a SRQL query is shown in Figure 3. A query is composed of the following clauses.

```
SELECT [FIRST | LAST | ( CLOSEST [LEFT | RIGHT | UP |
  DOWN] '(' < variable > ')' )] BOX | < property > [',' <
  property >...]
FROM <uri>
WHERE <boolean expression using spatial relations and
  the BOX variable>
HAVING <boolean expression on property values>
ORDER BY <property > | XPOS '(' UL | LR ')' | YPOS '('UL
   | LR ')' [ASC | DESC] [ ',' ... ]
WITH < variable > '=' [< uri >|<coordinates >|<sub−query
  >][',' < variable > '=' ...]
```

Fig. 3. Overall structure of a SRQL query

SELECT This mandatory clause defines the information to be extracted by the query. According to the spatial relation theory, the language allows to select *specific boxes* from the box model associated to the data, and return their graphical representation (e.g., a graphic snapshot), specified through the **BOX** keyword, and/or a comma separated list of their attributes. Moreover, the SRQL syntax allows to use several operators to refine the selection. Among these, the **FIRST** and **LAST** operators make the **SELECT** return only the first or the last extracted box, respectively, with respect to the ordering given by the **ORDER BY** clause, and the **CLOSEST [UP, DOWN, LEFT, RIGHT]** operators select only the extracted box that is closest to a given reference box, with respect to one of the four basic spatial directions.

FROM This mandatory clause contains the *URI* identifying the resource to be queried.

WHERE This clause contains a boolean expression that determines if a box has to be selected by the query. The expression is built by combining the spatial relations with logical operations. Relations can be applied to the keyword **BOX** (which represents the box being evaluated by the query) and to the *box variables* defined in the **WITH** clause. These variables can generally represent *set of boxes*. SRQL also supports the keywords **ANY** and **ALL** to compare single boxes with set of boxes, using the same semantics of SQL.
Finally, a further keyword **EACH** is provided to support more complex extractions, with the following meaning. Given the expression **BOX** rel **EACH** B, for each box $b \in B$, the **EACH** construct calculates the set of all the boxes a such that $a \, rel \, b$ holds. The overall evaluation allows to collect and return all these boxes a.

ORDER BY This clause defines the ordering of the results. The ordering key can be any box property (see the **HAVING** clause below) or the coordinates, specified by the keywords **XPOS** and **YPOS**, respectively, of the upper-left (**UL**) or the lower-right (**LR**) points of the box. The ordering can be **ASC**ending or **DESC**ending, and more than one ordering key can be specified.

WITH This clause is used to define some particular (sets of) boxes on the box model and assign them to variables to be used in the **WHERE** clause. The tool automatically defines the **START** variable to point to the upper left corner of the *user image*. Other boxes can be specified using their absolute coordinates or as the result of a nested query.

HAVING This clause contains a boolean expression using the common comparison operators applied to *property identifiers*. If this clause is specified, the expression is evaluated on each box selected by the **WHERE** clause, and only the ones satisfying both the **WHERE** and the **HAVING** expressions are returned by the query. In particular, when the box has a textual content (represented by the *textcontent* property), the **LIKE** operator can be used to match it against a regular expression.

As an example, a query written in SRQL syntax is shown in Figure 4. This query allows to extract all the text below the uppermost image in a web page.

```
SELECT BOX
FROM http://somesite.net
WHERE BOX DOWN image
HAVING textcontent <> ""
WITH immagine=
 SELECT FIRST BOX
 FROM http://somesite.net
 HAVING tag="img"
 ORDER BY YPOS(UL) ASC
```

Fig. 4. An example of SRQL query

Once a query has been written, it can be saved for later reuse, and analyzed/executed. Indeed, the system first performs some syntax checks on the query, possibly providing meaningful error messages that allow the user to improve the query definition, and then the query is executed on the box model.

After a successful query, the user image of the data is rendered in the application interface, the boxes selected by the query (if any) are highlighted, and the corresponding properties are appropriately returned. Finally, it is worth noting that the query results can be exported in an XML file (where the boxes, represented by the corresponding portion of the user image, are saved as separate files) to be exploited for further analysis and manipulations.

III. Case Studies

In this Section we show two applications of the SRQ tool on web pages and maps.

A. Web Pages

This case study shows how to formulate and test queries that are able to extract the book/CD cover images from *Amazon* search result web pages. A quick analysis of the visual layout of such pages allows to individuate some properties of each cover image:

- is always included between the page header and footer;
- appears as the closest image on the left of the corresponding (book/CD) textual description block;
- is top aligned with the (book/CD) title line;
- is on the right of the result reference number.

```
SELECT *
FROM http://www.amazon.com/....
WHERE BOX LEFT ANY description AND
 (BOX DOWN ANY pager AND BOX UP ANY pager)
 AND BOX RIGHT ANY resultNumber
HAVING nodeName = "img"
WITH description = SELECT BOX
 FROM local:parent
 HAVING textContent LIKE "Buy.*",
pager = SELECT BOX
 FROM local:parent
 HAVING textContent LIKE "Page:(\s*[0-9]+)+",
resultNumber = SELECT BOX
 FROM local:parent
 HAVING textContent LIKE "[0-9]+\."
```

Fig. 5. SRQL query for the Amazon case study

Therefore, a possible natural extraction query that strongly reflects these properties may be designed as follows: "return the boxes having image type, which are between the page header and footer, on the left of a description block and on the right of a number". In particular, the page header and footer are individuated through an appropriate regular expression that matches the result pager, whereas the description block is identified by a text beginning with the word "Buy". The corresponding SRQL query is shown in Figure 5. The same query also appears in the query editor box of Figure 6, together with some corresponding results.

Fig. 6. The SRQ tool main interface

```
SELECT CLOSEST RIGHT(number) *
FROM http://www.amazon.com/....
WHERE BOX horizontallyAligned EACH number
HAVING nodeName = "img"
WITH number = SELECT BOX
 FROM local:parent
 HAVING textContent LIKE "[0-9]+\."
```

(a)

```
SELECT *
FROM http://www.amazon.com/....
WHERE BOX LEFT titles AND BOX RIGHT numbers AND
 BOX DOWN search
HAVING nodeName = 'img'
WITH titles = SELECT BOX
 FROM local:parent
 HAVING class = 'productTitle',
numbers = SELECT BOX
 FROM local:parent
 HAVING class = 'number',
search = SELECT BOX
 FROM local:parent
 HAVING class='relatedSearch'
```

(b)

Fig. 7. Alternative SRQL queries for the Amazon case study

To improve the query formulation and accuracy, more expert users may exploit the full tool capabilities by integrating the query with visual attributes and structural information obtainable from the page source HTML. Indeed, alternative versions of the query in Figure 5 are shown in Figures 7(a) and 7(b). For example, the query in Figure 7(a) uses the **CLOSEST RIGHT** operator together with an user defined spatial relation *horizontallyAligned* (i.e., "not above and not below") to simplify the spatial constraints, whereas the query in Figure 7(b) uses the HTML *class* attribute to make the sub queries in the **WITH** clause more accurate.

Finally, in order to collect all the extracted images, it is possible to click on the Export button, saving both the images (in their source format) and the corresponding properties (in an XML file).

B. Maps

This case study develops a didactic application that could be used to teach geography to primary school students. In particular, the teacher uses the tool to make visual queries on the world map in order to interactively show geographic relationships and properties of the countries to the students.

To this aim, the SRQ tool is applied to a set of geospatial data described through the *ESRI Shapefile* [18] standard. In such maps, objects (e.g., country boundaries, rivers, interest areas like parks, etc.) are defined by complex shapes, possibly associated with peculiar semantic properties such as toponym, coordinates, population, etc. In particular, the world map used in the application has been taken from [19], where each map object refers to a specific country. Then, the box model generated by the SRQ tool for this application contains the bounding boxes of all the world countries, associated with the supplied semantic properties.

```
SELECT BOX
FROM file:/C:/worldShapeFile.shp
WHERE BOX DOWN refArea
HAVING regionID = 19 AND subregionID = 5 AND area
  >160000
ORDER BY area DESC
WITH refArea =
 SELECT BOX
 FROM local:parent
 HAVING fips = "VE"
```

Fig. 8. A sample SRQL query on the world map

As an example, a sample query returning the largest American countries that are on the south of a given reference country could be formulated as "*return the countries that are under the reference, having an area greater than a reasonable minimum (to say, 160000 sq mi), and order them by area*". Figure 8 shows the corresponding SRQL query, where Venezuela is the reference country, the *regionID* and *subregionID* semantic properties are used to restrict the query to the American countries, and the *fips* property is used to identify Venezuela in the subquery.

However, unlike the common geographic meaning of the word "South", this query would exclude Brazil from the results, since it is not completely under Venezuela, as formally required by the **DOWN** relation used in the query. To refine the results, we can define a proper **SOUTH** relation through the following simple SRDL statements:

```
LEFT_STRICT = LEFT AND (NOT UP) AND (NOT DOWN)
RIGHT_STRICT = RIGHT AND (NOT UP) AND (NOT DOWN)
SOUTH = (NOT UP) AND (NOT LEFT_STRICT) AND (NOT
  RIGHT_STRICT)
```

The execution of the query in Figure 8, where the relation **DOWN** is replaced by **SOUTH**, produces the results shown in Figure 9.

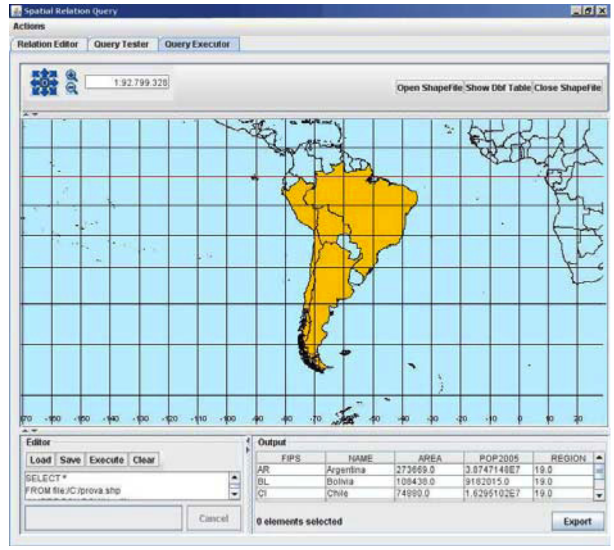


Fig. 9. SRQL query for the map case study

IV. Conclusions

In this paper we have presented a support environment for Information Extraction based on the visual appearance of the information. The underlying formal framework of spatial relations allows to give a sound theoretical foundation to this environment.

The SRQ tool is suitable to be used both by novel or experienced users. The former users may take advantage of the user-friendly interface and the natural query paradigm of the tool to accomplish their searches. The latter ones may fully exploit the integrated power of the tool and of the SRQL language by both appropriately customizing the spatial relations and extensively using visual, textual and structural constraints in the queries.

Currently, we are extending the tool by the addition of some refinements, including the filtering of invisible elements on the web pages, e.g., elements having the same color as the background, or overlapped elements, and the introduction of the third dimension (i.e., map layering) on geospatial data, through appropriate **ABOVE** and **BELOW** spatial relations. In the future, we also plan to work on the following two extended functionalities:

- *Relation definition by example*: the user will be able to define new spatial relations by sketching their *meaning* on a sample image. For example, once an image has been rendered, the user may simply draw the bounding boxes around some elements, mark one of them as a reference, and the tool will automatically generate the relation that holds between the reference box and the others.
- *Contour similarity predicates*: the user can sketch a contour and assign it to a variable. Then, the variable

can be recalled in a query to find boxes whose content matches the given contour.

References

[1] A. H. F. Laender, B. A. Ribeiro-Neto, A. S. da Silva, and J. S. Teixeira, "A brief survey of web data extraction tools," *SIGMOD Rec.*, vol. 31, no. 2, pp. 84–93, 2002.

[2] C. Chang and M. Girgis, "A survey of web information extraction systems," *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, no. 10, pp. 1411–1428, 2006.

[3] M. Lam and Z. Gong, "Web information extraction," in *Proceedings of the IEEE International Conference on Information Acquisition*, 2005.

[4] Y. Zhai and B. Liu, "Web data extraction based on partial tree alignment," in *WWW '05: Proceedings of the 14th international conference on World Wide Web*. New York, USA: ACM, 2005, pp. 76–85.

[5] G. Gottlob, C. Koch, R. Baumgartner, M. Herzog, and S. Flesca, "The LIXTO data extraction project - back and forth between theory and practice," in *Proceedings of the Symposium on Principles of Database Systems (PODS-04)*, 2004.

[6] A. H. F. Laender, B. Ribeiro-Neto, and A. S. da Silva, "DEByE - date extraction by example," *Data Knowl. Eng.*, vol. 40, no. 2, pp. 121–154, 2002.

[7] S. Soderland, "Learning information extraction rules for semi-structured and free text," *Machine Learning*, vol. 34, no. 1-3, pp. 233–272, 1999.

[8] H. Snoussi, L. Magnin, and J. Nie, "Towards an ontology-based web data extraction," in *BASeWEB Proceedings of the Fifteenth Canadian Conference on Artificial Intelligence AI 2002*, 2002.

[9] H. Zhao, W. Meng, Z. Wu, V. Raghavan, and C. Yu, "Fully automatic wrapper generation for search engines," in *WWW '05: Proceedings of the 14th international conference on World Wide Web*. New York, USA: ACM, 2005, pp. 66–75.

[10] W. Gatterbauer, P. Bohunsky, M. Herzog, B. Krüpl, and B. Pollak, "Towards domain-independent information extraction from web tables," in *WWW '07: Proceedings of the 16th international conference on World Wide Web*. New York, USA: ACM, 2007, pp. 71–80.

[11] V. Crescenzi and G. Mecca, "Grammars have exceptions," *Information Systems*, vol. 23, no. 8, pp. 539–565, 1998.

[12] J. Hammer, J. McHugh, and H. Garcia-Molina, "Semistructured data: The TSIMMIS experience," in *Proceedings of the First East-European Workshop on Advances in Databases and Information Systems*, 1997, pp. 1–8.

[13] Y. Aumann, R. Feldman, Y. Liberzon, B. Rosenfeld, and J. Schler, "Visual information extraction," *Knowl. Inf. Syst.*, vol. 10, no. 1, pp. 1–15, 2006.

[14] X. Gu, J. Chen, W. Ma, and G. Chen, "Visual based content understanding towards web adaptation," in *Proceedings of the Second International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems AH'02. LNCS*, vol. 2347. Springer-Verlag, 2002, pp. 164–173.

[15] Y. Yang and H. Zhang, "HTML page analysis based on visual cues," in *ICDAR '01: Proceedings of the Sixth International Conference on Document Analysis and Recognition*. Washington, DC, USA: IEEE Computer Society, 2001, pp. 859–864.

[16] G. Costagliola, A. De Lucia, S. Orefice, and G. Polese, "A classification framework to support the design of visual languages," *Journal of Visual Languages and Computing*, vol. 13, no. 6, pp. 573–600, 2002.

[17] ANTLR Parser Generator, "http://www.antlr.org/."

[18] Environmental Systems Research Institute, Inc., "Esri shapefile technical description," http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf, July 1998.

[19] B. Sandvik, "World borders dataset," http://thematicmapping.org/downloads/world_borders.php.

Rule-based Diagram Layout using Meta Models

Sonja Maier and Mark Minas
Universität der Bundeswehr München
{sonja.maier, mark.minas}@unibw.de

Abstract

This paper gives an overview of a drawing approach for the visualization of diagrams. The approach is tailored to editors for visual languages which support structured editing as well as free-hand editing. In this approach, the editor developer specifies layout behavior on top of a layout meta model (LMM). The LMM consists of two parts: one part models abstract syntax information of the diagram, and one part models concrete syntax information of the diagram that is relevant for layout computation. From the layout specification a drawing facility is generated. With the generated editor, the user may perform static and dynamic diagram drawing at any time. As running example, we choose UML class diagrams. Based on this example, we show the applicability of our approach to graph-based and other visual languages.

1 Introduction

When implementing an editor for a visual language, layout is a challenging task. On the one hand, the drawing approach should produce a good-looking result and should support the user. On the other hand, the layout specification should be very easy. In the following, we introduce an approach that aims at achieving these concurrent goals. The approach supports static and dynamic diagram drawing [1], which is triggered by user input. As a consequence, layout needs to be computed at runtime. The approach is also best suited for incremental layout, which starts with an initial layout and performs minor changes to improve it while still preserving the mental map of the original layout [12, 11]. Our approach succeeds in the context of editors that support structured editing as well as free-hand editing. Structured editors offer the user some operations that transform correct diagrams into (other) correct diagrams. Free-hand editors allow arranging diagram components on the screen without any restrictions. This high degree of freedom leads to the necessity of a very powerful layout engine. The specification of this engine is based on a layout meta model (LMM), instead of diagram-dependent internal representations. The use of meta models is a first step towards reusability of layout behavior.

Our approach was implemented and tested in DiaMeta [10], an editor generation framework. With this tool an editor developer may create an editor for a specific visual language, e.g., UML class diagrams. The result is an editor that can be utilized by the editor user to create UML class diagrams (see Figure 1). With DiaMeta, not only editors for graph-based visual languages like Petri nets, but also editors for other languages like GUI forms can be created. To all of these examples, our approach was applied successfully.

Section 2 describes the different layout modes available in our approach. In Section 3, a UML class diagram editor is introduced that serves as running example in this paper. Section 4 explains how to create a layout specification. The layout algorithm, which is applied after user interaction, is described in Section 5. Some implementation details are presented in Section 6. In Section 7 future work is summarized and the paper is concluded.

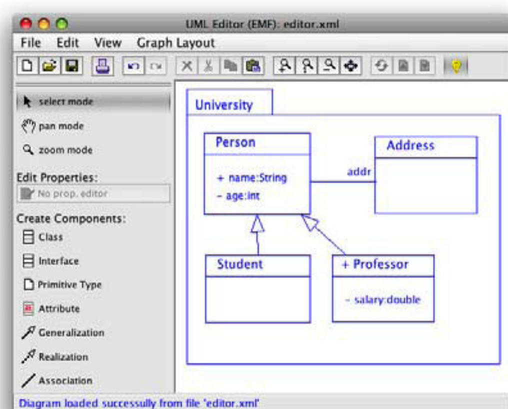
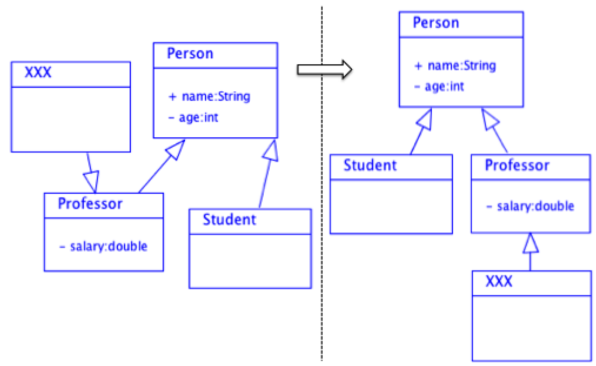


Figure 1. UML Class Diagram Editor

2 Layout Modes

First, we present an overview of the functionality our approach offers to the editor user. The layout engine may either be *called automatically* each time the diagram is changed by the user, or the layout engine may be *called explicitly*, e.g., by clicking a button. Furthermore, it may either be that a diagram is visualized the first time, or that a diagram has been visualized already. The layout engine may or may not take into account previous layout information. If no previous layout information is considered, it is called *static layout*. If previous layout information is taken into account, it is called *dynamic layout*. In the next paragraphs, we give an example of each of these two categories.

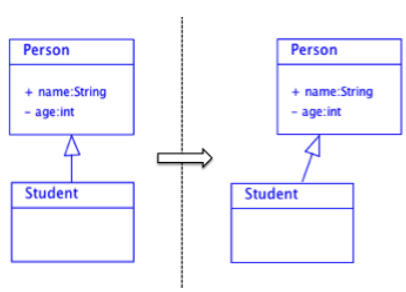


(a) (b)

Figure 2. Before (a) & After (b) Performing Static Layout

Static Layout In order to layout the UML class diagram as shown in Figure 2, the user may click the button "Graph Layout" (Figure 1). After clicking the button, the layout engine draws the diagram as shown, not considering previous layout information.

Dynamic Layout In Figure 3, the class `Person` is moved to the right. During movement, the arrow stays attached to the outlines of the components at any time. The layout engine draws the diagram as shown, considering pre-



(a) (b)

Figure 3. Before (a) & After (b) Moving the Class Person (Dynamic Layout)

vious layout information. In many tools, this kind of layout computation is already performed by the framework, and hence there is no need to explicitly define it. As our approach supports free-hand editing, this kind of behavior may also be defined.

3 Running Example - UML Class Diagrams

In general, we distinguish two categories of visual languages: graph-like and non-graph-like visual languages. Examples for the first category are Petri nets, mindmaps or business process models. Examples for the second category are Nassi-Shneiderman diagrams, VEX diagrams [5] or GUI forms.

In the following, we briefly revisit (a restricted version of) UML class diagrams, the visual language that will serve as running example. The language combines both, the concepts of a graph-based and a non-graph-based visual language. Classes together with generalizations and (or) associations form a graph, whereas packages, together with classes and attributes (also called properties) form a hierarchical visual language. In Figure 4, a UML class diagram, which was created with the UML class diagram editor, is shown. In the tool, the concrete syntax of the visual language is represented by a graph. In the graph, the nodes (component nodes) that are visualized as rectangles represent the components contained in the diagram: one package, one class and two attributes. The edges (relationship edges) `nextProperty` , `containedProperty` and `containedClassifier` that are visualized as arrows represent geometric dependencies between attribute and attribute, class and attribute, and package and class respectively. These edges are connected with the component nodes via nodes visualized as circles (attachment areas) and edges (attachment edges) visualized as lines.

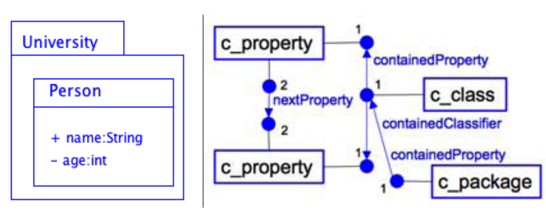


Figure 4. Class Diagram & Graph Model

4 Layout Specification

In the following, we describe, how a layout specification may be created. The layout specification consists of three parts: The Layout Meta Model (LMM), a set of layout rules and the application control.

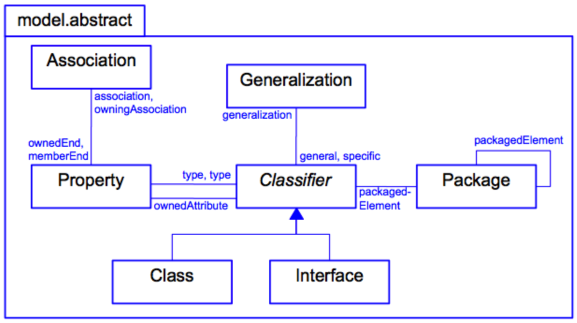


Figure 5. Abstract Syntax Meta Model

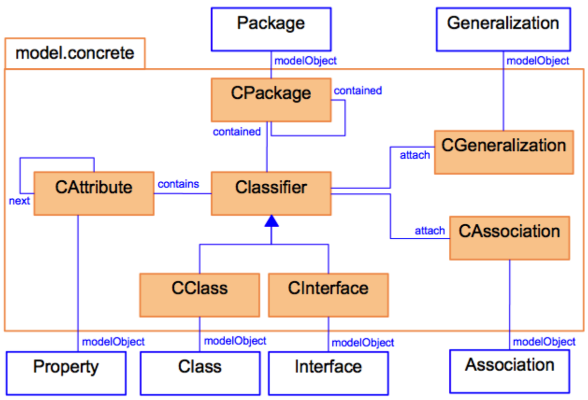


Figure 6. Concrete Syntax Meta Model

4.1 Layout Meta Model

For our approach, information of the abstract and concrete syntax is required. It may be accesed via the LMM: On the concrete syntax level, e.g., a class is represented by a rectangle with a certain size, or attributes are visualized in a certain order. On the abstract syntax level, for instance, information about nesting of packages is available.

In general, the LMM consists of two parts: The *abstract syntax meta model* (Figure 5) and the *concrete syntax meta model* (Figure 6). These two parts are connected via `modelObject` links. In our example, the abstract syntax meta model is similar to the UML2 Ecore specification [15]. The concrete syntax meta model contains geometric information that is needed for layout computation.

The example diagram shown in Figure 7 consists of the packages `package1` (p1) and `package2` (p2), the class `Class1` (c1) and the attributes `attribute1` (a1) and `attribute2` (a2). The corresponding instance of the LMM is shown in Figure 7 (bottom).[1] In the abstract syntax model (Figure 7, right), there is no link between `c1` and `p1`. In the concrete syntax model (Figure 7, left), there is an additional link between the attributes `a1` and `a2`. This link gives us information about the order of the attributes.

[1]To increase readability, modelObject links have been removed in the diagram.

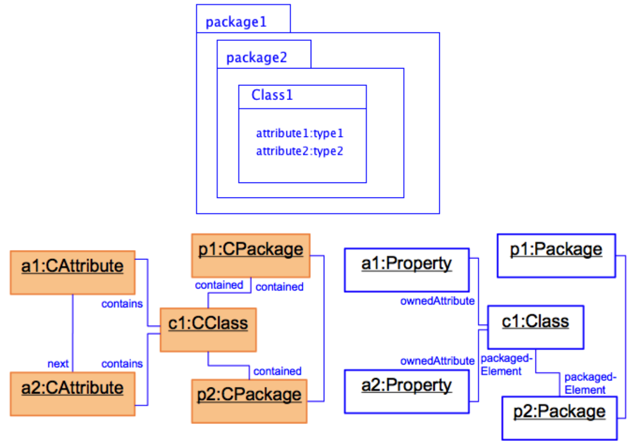
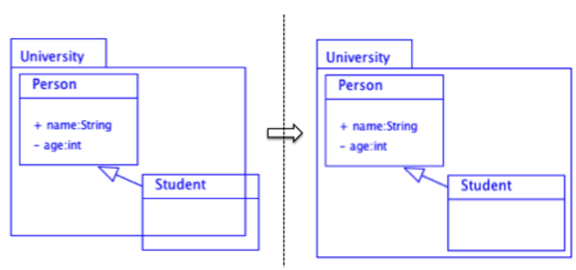


Figure 7. Example Diagram & Instance of LMM

4.2 Layout Rules

Layout rules can be defined based on the LMM. They specify local attribute changes of components, and are coordinated by the application control. In the following, we describe two layout rules that are used for the UML class diagram editor.

Dynamic Layout To define a layout rule usable for dynamic layout computation, we need to proceed as follows: the developer has to provide a pattern, an optional condition and an action. The pattern identifies a component, which has been changed previously (either by the user or by the layout engine), together with the local context needed for layout computation. The condition describes the circumstances, in which an action is applied. The action defines the attribute changes that are performed if the condition is fulfilled.



(a) (b)

Figure 8. Before (a) & After (b) Resizing the Package University

In Figure 8, a UML class diagram before and after applying a layout rule that resizes the package `University` is shown. Figure 9 shows the corresponding layout rule. The pattern consists of a class and a package. The class `cc1` is the component that has been changed by the user. In the example, the condition checks, whether the class sticks out to the right of, or to the bottom of, the package. If this is not the case, the action updates the attributes `x2` and `y2` of the package `cp1`, and hence, the package is enlarged: If it sticks out to the right, the width is enlarged, if it sticks out to the bottom, the height is enlarged.[2] Here, the point `(x1,y1)` is the top left corner of the component, and `(x2,y2)` is the bottom right corner respectively.

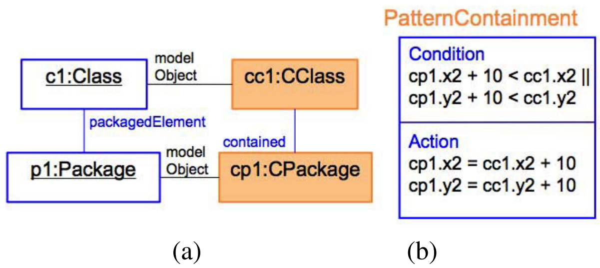


Figure 9. Pattern (a), Condition & Action (b)

Static Layout Defining a layout rule usable for static layout is a special case of defining a layout rule usable for dynamic layout: all components are marked as changed. Then the layout is computed as usual, without having initial values for attributes available.

Visual Editor To allow for a more intuitive description, we use a visual language for layout rule definition, which is based on the concrete syntax meta model. The language was introduced in [8]. The use of a visual language is possible as layout rules are often defined solely on the concrete syntax meta model. This way it is possible to specify them visually in the application domain. A generic editor for specifying these rules is provided. This language-specific editor is based on a diagram editor for the specified visual language, which does not support layout yet. In Figure 10, a screenshot of such an editor for UML class diagrams is shown. In order to create a layout rule, the editor developer draws a pattern, and enters the pattern name, a condition and an action in the middle of the editor (from top to bottom). For specifying the pattern, he may use the components available on the left side of the editor. In the example, he uses the components `Class` and `Generalization`. For specifying conditions and actions, he may access the attributes shown on the right side of the editor.

[2] A similar rule exists for the case that the class sticks out to the left of, or to the top of, the package.

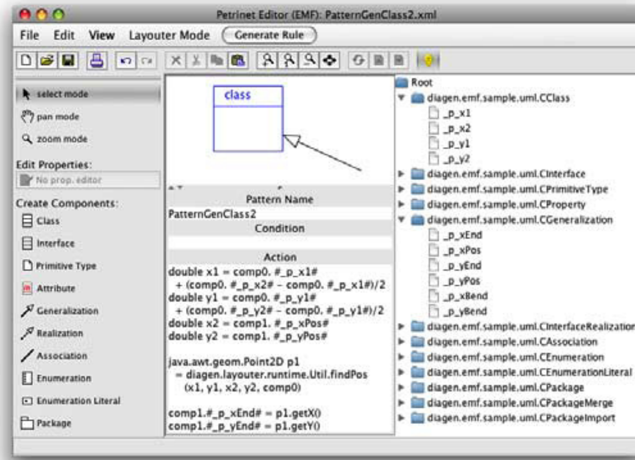


Figure 10. Layout Editor

In the example, he uses the attributes `x1`, `y1`, `x2`, `y2` of the component `class`, and `xPos`, `yPos`, `xEnd` and `yEnd` of the component `generalization`.

4.3 Definition of Application Control

Besides LMM and layout rules, the application control needs to be defined. The application control coordinates the application of layout rules (called strategies) and the interplay between different layout rules (called interplay). When defining the interplay, it is also possible to plugin an external layout algorithm like the *Sugiyama Algorithm* [14]. The definition of strategies and the interplay is based on the LMM of the diagram language.

Strategies In the following, we describe three strategies that are used for the UML class diagram editor.

If a class, e.g., `Class1` of the diagram in Figure 7, is moved or resized, the strategy dictates the following behavior: follow the link `packagedElement` (cf. Figure 7) and retrieve the surrounding package `package2`. An associated layout rule is applied to this package, which may move or resize the package. If the package has been changed, the strategy dictates the following behavior: follow the link `packagedElement`, and retrieve the surrounding package `package1`. To this package, another associated layout rule is applied. The specification of the two strategies ("a class is changed" or "a package is changed") is based on the *abstract syntax meta model*.

If an attribute is moved, e.g., the attribute `attribute1` of the diagram in Figure 7, the (x,y) position of the other attributes, e.g., of the attribute `attribute2`, needs to be updated. Here, an application strategy ("attribute is moved") is required, which is based on the *concrete syntax meta model*.

When moving a class or interface, generalizations and associations that are connected to this class are updated. Here, the application control is quite simple, as the change of generalizations or associations has no other effect. This means, that *no strategy* needs to be defined.

We have not yet introduced a language for specifying strategies. Plain Java code must be programmed instead. However we plan to introduce a language based on PROGRES or Fujaba [4].

Interplay After choosing certain strategies, it needs to be defined how these strategies interact with each other. Here, many variants are conceivable. The easiest interplay that may be defined is applying strategies one after another. For the UML class diagram editor, the following variant is used: first, a strategy that moves packages, classes and interfaces to the correct position is applied. Here, it is possible to position the classes using a standard graph drawing algorithm like the *Sugiyama Algorithm*. Afterwards, a strategy that relocates attributes is executed. In the last phase, a strategy that updates associations and generalizations is applied.

5 Layout Execution

If the layout engine is called, roughly speaking, an instance of the LMM is created from the diagram, and several layout rules are applied, following the strategies and the interplay described in the application control.

The algorithm is based on ideas described in previous work. In the extended approach presented here, layout rules as well as the application control are defined on top of the LMM, instead of the internal graph structure of a specific editor [8, 9], or on the abstract syntax meta model [7].

5.1 Layout Meta Model

After changing the diagram, first of all, an instance of the LMM is created by the diagram editor. This model is derived from the updated graph representation of the diagram (an example is shown in Figure 4). Additional information available in the abstract syntax model is retrieved through syntax analysis.

5.2 Application Control

Generally, the interplay defines the order in which strategies are applied. Each strategy describes to what part of the diagram a rule is applied.

Interplay First, the interplay decides on the sequence of applied strategies. In our example, packages, classes and interfaces are moved to the right position, first. Afterwards, attributes are relocated, and then, associations and generalizations are moved to the right position.

Strategies Each strategy determines in which order layout rules are applied to different parts of the diagram. For instance, internal packages are moved and resized first, and surrounding packages are updated afterwards.

5.3 Layout Rule

For a layout rule, one diagram component is given as input. This is usually a component, the user changed, or a component the layout engine changed. E.g., for the rule ”resize packages” described earlier, this is the class. Then other components that are necessary for the layout rule are identified. For the example layout rule, this is the package. Afterwards, if a certain condition is fulfilled, the corresponding action is applied. For the example layout rule, the condition checks if the ”package is big enough”. If this is not the case, the action executes ”enlarge package”. This means that layout attributes are changed. In our example, the attribute `x2` (or `y2`) is changed, and hence the package is enlarged.

6 Technical Realization

The approach has been implemented in DiaMeta [10], an editor generation framework. The editor generation framework DiaMeta provides an environment for rapidly developing diagram editors based on meta modeling. Figure 1 shows a DiaMeta editor for UML class diagrams. Each DiaMeta editor is based on the same editor architecture, which is adjusted to the specific diagram language.

The DiaMeta environment consists of a framework and the DiaMeta Designer. The framework is basically a collec-

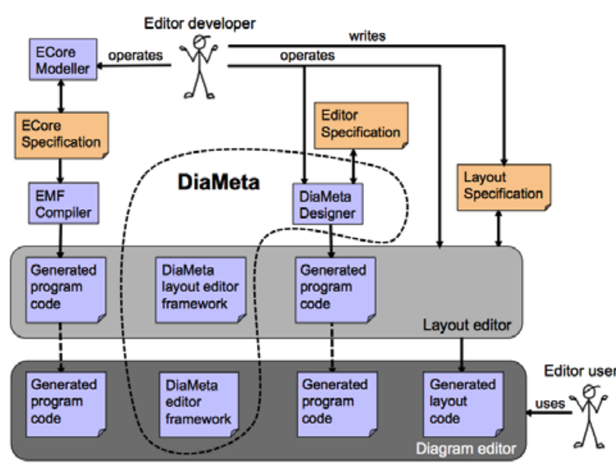


Figure 11. DiaMeta Framework

tion of Java classes and provides the dynamic editor functionality, necessary for editing and analyzing diagrams. In order to create an editor for a specific diagram language, the editor developer has to enter two specifications: The abstract syntax meta model (EMF model [3]), and the editor specification. The language's abstract syntax meta model is defined by a class diagram. The EMF compiler is used to create Java code that represents this model. The editor developer uses the DiaMeta Designer for specifying the concrete syntax and the visual appearance of diagram components, e.g., classes (UML class diagram) are drawn as rectangles. The DiaMeta Designer generates Java code from this specification.

If automatic layout is designated, the editor developer additionally has to provide a layout specification, meaning the LMM, some layout rules, and the application control. The LMM is an extension of the abstract syntax meta model. The EMF compiler is used to create Java code that represents this extended model. The rest of the layout specification, meaning the rules as well as the application control, is based on this LMM.

Some parts of the layout specification are defined via the layout editor (see Figure 10), an extended editor that does not support automatic layout, yet: The Java code created by the EMF compiler, the Java code generated by the DiaMeta Designer, and the layout editor framework, implement this layout editor for the specified diagram language. The Java code created by the EMF compiler, the DiaMeta Designer, and from the layout specification as well as the editor framework, implement a diagram editor for the diagram language, which supports automatic layout.

7 Future Work & Conclusions

In this paper, we had given an overview of a drawing approach for the visualization of diagrams. The approach is tailored to an editor for visual languages, which supports structured editing as well as free-hand editing. The editor developer specifies layout behavior. From this specification a drawing facility is generated. With the generated editor, the user may perform static and dynamic diagram drawing. The approach was tested for many visual languages, like Petri nets or GUI forms. In this paper, we chose the visual language UML class diagrams as running example.

There exist several editor generator frameworks, but none of them allows for the generation of a (powerful) layout engine. For drawing graphs, some related approaches exist: a general discussion can be found in [1] and a tool that focusses on dynamic graph drawing is Dunnart, which uses a constraint-based drawing approach [6].

In our approach, we only consider correct parts of the diagram. This leads to some difficulties during diagram creation. Besides, an open issue is the reusability of layout rules and the application control. At the moment, they are rewritten each time they are needed. To cope with this challenge, we are going to apply the concepts used in [13]. Up to now, we primarily focus on achieving "nice" results. In future, we also would like to take care of easy diagram manipulation. In this context, some user studies will need to be performed. To evaluate the "degree of mental map preservation", some metrics will be helpful [2].

References

[1] J. Branke. Dynamic graph drawing. In *Drawing graphs: methods and models*, pages 228–246. Springer-Verlag, 2001.

[2] S. S. Bridgeman and R. Tamassia. Difference Metrics for Interactive Orthogonal Graph Drawing Algorithms. In *Proc. of the 6th Intl. Symposium on Graph Drawing (GD '98)*. Springer-Verlag, 1998.

[3] F. Budinsky, S. A. Brodsky, and E. Merks. *Eclipse Modeling Framework*. Pearson Education, 2003.

[4] C. Fuss et al. The Jury is still out: A Comparison of AGG, Fujaba, and PROGRES. In *Proc. of the 6th Intl. Workshop GT-VMT'07*, volume 6 of *ECEASST*, 2007.

[5] W. Citrin, R. Hall, and B. Zorn. Programming with visual expressions. In *Proc. of the 11th Intl. IEEE Symposium on Visual Languages (VL '95)*. IEEE Computer Society, 1995.

[6] T. Dwyer, K. Marriott, and M. Wybrow. Dunnart: A Constraint-Based Network Diagram Authoring Tool. 2008.

[7] S. Maier, S. Mazanek, and M. Minas. Layout Specification on the Concrete and Abstract Syntax Level of a Diagram Language. In *Proc. of the 2nd Intl. Workshop on Layout of (Software) Engineering Diagrams (LED 2008)*, volume 13 of *ECEASST*, 2008.

[8] S. Maier, S. Mazanek, and M. Minas. Visual Specification of Layout. In *Proc. of the 16th Intl. Symposium on Graph Drawing (GD '08)*. Springer-Verlag, 2008.

[9] S. Maier and M. Minas. Specication of a Drawing Facility for Diagram Editors. In *Proc. of HCI International 2009 (HCII 2009)*, Springer-Verlag, 2009.

[10] M. Minas. Generating Meta-Model-Based Freehand Editors. In *Proc. of the 3rd Intl. Workshop on Graph Based Tools*, volume 1 of *ECEASST*, 2006.

[11] K. Misue, P. Eades, W. Lai, and K. Sugiyama. Layout Adjustment and the Mental Map. *Journal of Visual Languages & Computing*, 6:183–210, 1995.

[12] H. C. Purchase and A. Samra. Extremes Are Better: Investigating Mental Map Preservation in Dynamic Graphs. In *Proc. of Diagrams '08*. Springer-Verlag, 2008.

[13] C. Schmidt, B. Cramer, and U. Kastens. Usability Evaluation of a System for Implementation of Visual Languages. In *Proc. of the IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC '07)*, 2007.

[14] I. G. Tollis, G. Di Battista, P. Eades, and R. Tamassia. *Graph Drawing: Algorithms for the Visualization of Graphs*. Prentice Hall, 1998.

[15] UML2: EMF-based implementation of the UML 2.x OMG specification. http://www.eclipse.org/uml2/, 2008.

Chorem Maps: towards a Legendless Cartography?

Robert Laurini[1], Françoise Raffort[1], Monica Sebillo[2], Genoveffa Tortora[2], Giuliana Vitiello[2]
[1]*LIRIS, INSA de Lyon, F-69621 Villeurbanne*
Robert.Laurini@insa-lyon.fr
[2]*DMI, Università di Salerno, Italy*
{msebillo, tortora, vitiello}@unisa.it

Abstract

There is a paradox between usual cartography for which legends are mandatory, and visual languages for which some textual explanations will represent a failure. In this paper, we address this paradox by presenting various points of view. Then, we will argue that chorems as schematized representations of territories issued from both cartography and visual languages can bring a solution in this paradox. To conclude, we claim that no legends are necessary when chorem maps are well designed.

1. Introduction

As it is common to present maps with legends, visual languages usually do not go with legends. They have to visually convey concepts preventing any kind of ambiguities.

On the other hand, chorems as schematized representations of territories derive both from cartography and visual languages. Then the question is: do we have to make legends for chorems? In this paper, we will address this difficult question by developing some arguments.

As a preliminary, let us remind that the word “legend” comes from Latin meaning “things to be read”.

This paper is organized as follows. After having described the scope of legends in cartography, visual languages will be discussed. Then, the special case of chorems will be examined. In the final remarks, we will argue that well-defined chorems do not need legends.

2. Cartography

Maps are used from millennia to describe territories. Among them, one of the oldest is the so-called Babylonian "map of the world" located in the British Museum, London (See Fig.1). Apparently, it is a Neo-Babylonian (Persian Period, circa 500 BCE) copy of an original dating to the Sargonid Period, circa late eighth or seventh century BCE. It is likely that maps older than this were designed and in use.

Figure 1. Babylon Map of the world (circa 500 BC) [14].

During the same period, also Hecataeus of Miletus (c.550BC-c.490BC), a Greek geographer and researcher, designed a world map. Furthermore, we know that maps of Roman roads existed. According to [15], the so-called map of Peutinger can be dated to the twelfth or thirteenth century, but it is clear that it is a copy of a much older original. It is probably identical to the world map that was prepared by Marcus Vipsanius Agrippa (64-12BC), a personal friend of the emperor Augustus [18]. After his death, this map was carved into marble and set up in Roma.

Anyway, all those maps were made without legends.

Now, it is not standard practice for cartographers to create maps without legends. We do not know who the first person to give this order was. Throughout history, legends can be found, for instance in the French Cassini map from the XVIIth century.

Let us remind that each map features a legend, or a key, which explains the meaning of each map symbol, sometimes named symbology. The selection of symbols, their shape and colours are usually made by the cartographer. His skill or expertise is to select the symbols which are the most understandable.

A map legend distinguishes particular types of lines, showing that they depict rivers, roads, railroads, or political boundaries. A legend also indicates how cities are depicted with squares or dots, according to their sizes. It illustrates the use of colour to show different elevations, political entities or different attributes of objects. A map legend also provides some or all of the following important information about a map: its title, scale, projection, cartographer, publisher, date of compilation, and date of printing.

As a conclusion, let us say that for general purpose maps, the legend is obvious whereas for thematic maps, due to author's choice, a legend is necessary in order to disambiguate the meaning of the depicted objects and phenomena. However, some textual labels could be necessary, such as the name of countries.

From a different point of view, do not forget that during several centuries, features were designed as we know them: the best example can be the Egyptian Nebamon Garden [16] from Thebes dated 14th century BC. This is the only way to avoid ambiguities. In the Renaissance, after Brunelleschi (1415) and Alberti (1435) [17] found perspective rules, features were designed as we see them: the main consequence was the existence of ambiguities, so the necessity to disambiguate some situations.

3. Visual Languages

The widely recognized value of icons, diagrams and other graphical notations as a means for representation have caused the development of several different languages and systems based on a fully visual approach, known as visual languages and environments, respectively.

In literature several taxonomies have been proposed to classify them based on different dimensions, which have given rise to different classification schemes. Among the others in [12] Shu categorizes visual languages in terms of the *level of the language*, its *scope or generality*, and the *extent of visual expressions employed*, whereas in [2] Chang classifies visual languages on the basis of the kind of activity they support, i.e., *visual interaction*, *visual programming*, *visual information processing* or *iconic visual information processing*. Finally, in [8] visual languages and commercial products are classified by focusing the attention on the kind of visual representation adopted by each language and on the (possible) combined use of textual and visual information. In particular, the authors propose a scheme where *pure* visual languages are those languages which make no use of textual notation. The semantics of the language is based on a purely graphical set of transformation rules whose comprehension require no knowledge of the underlying textual language. On the contrary *hybrid* visual languages exploit a combination of textual and visual representations. The authors also note that text can be more suited for distinguishing between elements that are of the same kind and for expressing concepts that are inherently textual, such as algebraic formulas.

A common visual representation that characterizes a well-experimented subclass of visual languages is the *iconic* representation. An iconic sentence is a spatial arrangement of icons, which are pictorial representations of conceptual entities and operations [3]. Each icon is an object with the dual representation of a physical part (the image) and a logical part (the meaning). Figure 2 depicts the underlying structure of an icon. This aspect makes icons able to express computation and distinguishes them from the pictures in a drawing program, which do not have such a semantics.

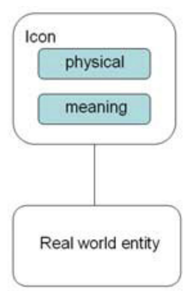


Figure 2. The icon structure

Figure 3 shows an icon representing Italian regions, among which the Campania region is highlighted. This example shows the capability of icons to convey logical information about a geographic area, by distinguishing it from its surroundings.

One of the most appealing features of the class of iconic visual languages is the fact that it has a well-established theoretical background, derived from the deep investigation following Chang's early work [3].

In agreement with such a formalization, two kinds of icons are considered, namely *process icons* and *object icons*.

The former express computations, the latter can be further subdivided into *elementary* object icons and *composite* object icons. Elementary object icons identify primitive objects in the language, whereas composite object icons identify objects formed by combining on the screen elementary object icons by using spatial operators, such as *hor* (for horizontal composition), *ver* (for vertical composition), *ovl* (for overlay), and *con* (for connect). The design methodology of iconic languages is based on a formal icon specification and the icon algebra. The former handles the static aspect of an iconic system. The latter algebra handles the dynamic aspect of an iconic system.

Figure 3. An icon representing the Campania region

In the same period, sharing a similar point of view and exploiting Chang's basic definition, Tortora defined an *iconic programming language* as a triple <*ID*, *G*, *B*>, where *ID* is the icon dictionary, *G* is a grammar, and *B* is a domain-specific knowledge base [13]. The icon algebra underlying the dynamics of the programming languages is specified by a grammar whose rules allow to construct composite object icons from elementary icons by using spatial arrangement operators, one can exploit a grammar $G = (N, T, OP, S, P)$, where N is the set of non terminals, T is the set of terminals denoting icons, OP is the set of spatial relational operators, S is the start symbol, and P is the set of productions whose right hand side is an expression involving relational operators. The knowledge base B contains domain-specific information necessary for constructing the meaning of a given visual sentence. It contains information regarding event names, conceptual relations, names of resulting objects, and references to the resulting objects.

The design methodology of iconic languages is based on a formal icon specification and the icon algebra. The former handles the static aspect of an iconic system. The latter algebra handles the dynamic aspect of an iconic system. The iconic indexing technique, together with the concept of icon similarity, allows the indexing and comparison of generalized icons.

Many domain applications may benefit from iconic programming languages, such as image processing, image database design, document editing, robotic manipulation, and VLSI design. Moreover, the concept of generalized icons and icon algebra also provide a unified framework for the theory of pattern recognition.

4. Chorem Maps

From 2006 three research institutions have put in action an international collaboration in the field of the geographic information, namely Institut National des Sciences Appliquées (INSA) of Lyon (France), Dipartimento di Matematica e Informatica (DMI), University of Salerno (Italy) and Tecnólogico de Monterrey, Campus de Puebla (Mexico). The goal of the involved researchers is to provide decision makers with advanced tools of geovisualization and geovisual analytics, capable of semantically and visually reasoning about quantitative, qualitative and cognitive aspects of a domain of interest.

The key element on which research activities have been based is represented by the *chorem* concept. According to its original meaning, a chorem is a schematic territory representation, which eliminates details not necessary to the map comprehension [1]. They were introduced in 1980 by Pr. Roger Brunet, a French geographer, who proposed a basic vocabulary (in French), made up of 28 elementary chorems, as shown in Figure 4. In particular, he stated that seven shapes are sufficient to describe models representing chorems, namely:

- *point*, meant to represent places, spots, poles;
- *line* meant to represent contacts, breaking, ties, borders, relations;
- *area*, meant to represent extensions and shapes;
- *flow*, meant to represent dynamics, movements, symmetry, intensity;
- *passage*, meant to represent bridges, crossing, bifurcations, tunnels;
- *polarization*, meant to represent focusing, dynamics;
- *gradient*, meant to represent dissymmetry, attraction, repulsion.

Since then, different proposals of chorems and several uses of the underlying concept can be found in literature, each sharing the common goal to represent spatio-temporal phenomena referring to a territory. In [7] a survey of the most significant proposals about chorems can be found. Moreover, in [10] more information can be obtained regarding a controversy about the philosophy behind chorems and their use.

Provided the primary capability of a chorem to represent a phenomenon through an elementary spatial configuration, the pioneering work by Laurini *et al.* has investigated the potentialities of chorems and listed different roles that they may play in supporting expert users' daily activities [9]. The authors state that chorems can be used to represent geographic knowledge, to visually summarize database contents and finally to underlie the creation of a novel progressive entry system for geographic databases, by following Ben Shneiderman's *mantra* for designing human interfaces [11].

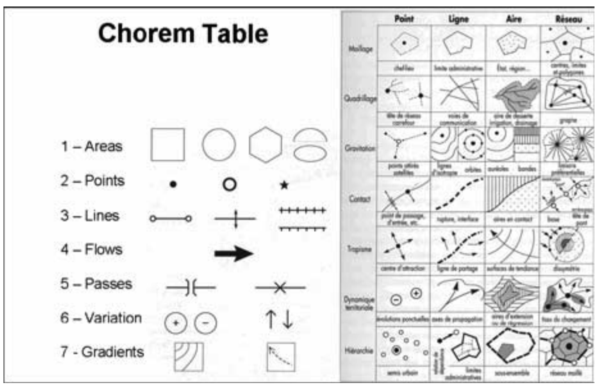


Figure 4. Table of Brunet's chorems.

Significant results have been obtained along this line, which support the initial proposal aiming at formally enriching the chorem meaning, in order to reach a new cartographic solution capable to support analysts' work in representing dynamics, movements and changes that underlie possible problems [4], [5], [6]. In particular, as for the chorem capability to visually summarize database content, in [6] the results concerning this aspect are presented. The concept of chorem has been adopted and formally extended, and a construction process of maps of chorems is specified by showing the tasks for both extracting relevant information from a large amount of data, and representing them in a form suitable for consumption by the ultimate users.

Figure 5 illustrates a map obtained through the proposed prototype referring to the Italian migrating population in 2000. It represents an immediate synthesis of data of interest by associating them with schematic visual notations, namely the Italian territory and migrating flows concerning it.

Suppose that in place of the symbol of a human, there is a truck, it is immediately understandable that arrows will represent flows of goods. In contrast, if arrows are not decorated with symbols, the meaning of flows is undetermined, because no meaning can be implicitly assigned the context.

We claim that in Figure 5, no legends are necessary.

The forthcoming research we are planning is meant to demonstrate that a set of chorem maps deriving from a unique spatial database can be used to both catch a thematic global view of a territory and its phenomena, and investigate a single phenomenon by accessing data characterizing it. Indeed, based on previous results, we aim to enhance the role of a chorem map with an added value deriving from the synergic combination of analysis techniques, visual query languages and human capability to reason and decide. In this manner, the visual notation used to display chorems will be always productive, and it may be exploited also to go back data which contributed to the phenomena under investigation.

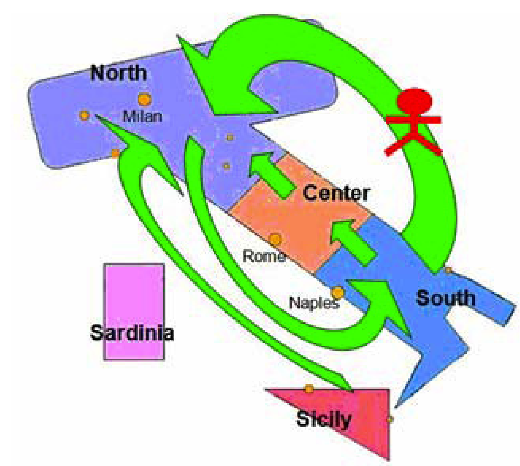


Figure 5. An example of a chorem map: Italian migrating people in 2000.

5. Final Remarks

In the field of geographic information, there exist many applications which require an immediate and intuitive cartographic support. This is due to the heterogeneous expertise that the ultimate users coming from specific application domains may have, and that may affect benefits resulting from their activities.

The awareness of this requirement has led to alternative approaches in the contest of map projections and cartographic generalization which have been focused on conveying rough ideas about the spatial arrangement of objects and the presence of dynamics, disregarding geometric aspects, such as scale and dimensions.

In this paper we have initially summarized motivations underlying an international collaboration among three research institutions. Then, we have briefly highlighted some aspects of the contribution that institutions are giving to the project, which result strictly related to the experience acquired in specific research fields. In particular, in this paper we have investigated a general-purpose matter aiming at stimulating the involved communities towards a constructive comparison in order to understand the real contribution that visual languages are giving to the modern cartography. In particular, our goal is twofold. We aim to both determine aspects that visual languages and cartography share in terms of visual notations and semantics, and recognize properties of visual language fundamentals that can (visually) enhance the expressiveness of chorem maps. In particular, we refer to the theory of icons, or the theory of dual representations of objects and its semantics, seen as a unified methodology for visual language design as well as icon-oriented system design.

To conclude this paper, let us say that well-defined chorems must follow well-adapted conventions and habits for representing symbols and selecting colours. So, we claim that well-defined chorem maps do not need legend.

6. References

[1] Brunet R. La carte-modèle et les chorèmes. Mappemonde 1986; 4: 4–6.

[2] S.K. Chang, G. Tortora, A. Guercio and Bing Yu "Icon Purity - Toward a Formal Theory of Icons", International Journal of Pattern Recognition and Artificial Intelligence, Vol. 1, No. 3&4, 1987, 377-392.

[3] S. K. Chang, Ed. *Visual Languages and Visual Programming*, Plenum Publishing Corporation, New York, 1990.

[4] V. Del Fatto, R. Laurini, K. Lopez, R. Loreto, F. Milleret-Raffort, M. Sebillo, D. Sol-Martinez, G. Vitiello "Potentialities of Chorems as visual summaries of spatial databases contents", Proceedings of the *9th International Conference on Visual Information and Information Systems VISUAL2007*, Shanghai (CINA), 28 – 29 May 2007. G. Qiu et al. (Eds.): VISUAL 2007, LNCS 4781, pp. 537–548, 2007. © Springer-Verlag Berlin Heidelberg 2007.

[5] V. Del Fatto, R. Laurini, K. Lopez, M. Sebillo, G. Tortora, M. Tucci, G. Vitiello, "Chorem Editing - Visual Summary of Spatial Database Content", in Proceedings of the 13rd International Conference on Distributed Multimedia Systems - Workshop on Visual Languages and Computing (VLC2007)", San Francisco USA, 6-8 Sept. 2007, pp.256-259.

[6] V. Del Fatto, R. Laurini, K. Lopez, M. Sebillo, G.Vitiello, "A chorem-based approach for visually synthesizing complex phenomena", in *Journal of Information Visualization*, vol. 7, n. 3, 2008, pp. 253-264. G. Andrienko, N. Andrienko, J. Dykes, S. Fabrikant and M. Wachowicz (Eds). Palgrave Macmillan Journals.

[7] V. Del Fatto, "Visual Summaries of Geographic Databases by Chorems", PhD jointly-awarded PhD in Computer Science, INSA Lyon, France and University of Salerno, Italy. April 7, 2009.

[8] F. Ferrucci, G. Tortora, G. Vitiello, "Visual Programming", 2nd edition of *Encyclopedia of Software Engineering*, J.J. Marciniak (Ed.), John Wiley and Sons, 2001, ISBN: 0-471-37737-6

[9] R. Laurini, F. Milleret-Raffort, and K. Lopez "A Primer of Geographic Databases Based on Chorems" R. Meersman, Z. Tari, P. Herrero et al. (Eds.): *OTM Workshops 2006*, LNCS 4278, pp. 1693 – 1702, 2006. © Springer-Verlag Berlin Heidelberg 2006.

[10] R. Laurini, M. Sebillo, G. Vitiello, D. Sol-Martinez, F. Raffort "Computer-generated Visual Summaries of Spatial Databases: Chorems or not Chorems?" *S.A.P.I.EN.S*, 2.2 | 2009, [Online], Online since 03 June 2009. URL : http://sapiens.revues.org/index795.html. Connection on 29 juin 2009.

[11] Shneiderman B. "Designing the User Interface." Third edition. Addison-Wesley, Publishing Company. 1997, 600pp.

[12] N. C. Shu, "Visual Programming Languages: A Perspective and a Dimensional Analysis," in S.K. Chang, T. Ichikawa, and P.A. Ligomenides, editors, *Visual Languages*, Plenum Publishing Corporation, New York, 11-34, 1986.

[13] G. Tortora "Structure and Interpretation of Visual Languages", in *Visual Languages and Visual Programming*, ed. S. -K. Chang, Plenum Press, New York, 3-30 (1990).

[14] Retrieved June 8, 2009 from http://members.optusnet.com.au/gtosiris/Babylonian-World-Map.jpg

[15] Retrieved June 8, 2009 from http://www.livius.org/pen-pg/peutinger/map.html

[16] In Wikipedia *The Free Encyclopedia*. Retrieved June 8, 2009 from http://upload.wikimedia.org/wikipedia/commons/thumb/4/4b/Pond_in_a_garden.jpg/400px-Pond_in_a_garden.jpg

[17] In Wikipedia *The Free Encyclopedia*. Retrieved June 8, 2009 from http://en.wikipedia.org/wiki/Perspective_(graphical)

[18] In Wikipedia *The Free Encyclopedia*. Retrieved June 8, 2009 from http://en.wikipedia.org/wiki/Marcus_Vipsanius_Agrippa.

Preserving the Hand-drawn Appearance of Graphs

Beryl Plimmer[1], Helen Purchase[2], Hong Yul Yang[1], Laura Laycock[1]

[1]Department of Computer Science
University of Auckland
Private Bag 92019
Auckland, New Zealand

[2] Computing Science Department
University of Glasgow
17 Lilybank Gardens
Glasgow, Scotland

ABSTRACT

When using a sketching tool to draw graphs, the edges need to appear hand-drawn. This is particularly the case after edges have been repositioned – if the action of moving a node results in its edges appearing as straight lines, the graph drawing will not retain its informal, hand-drawn appearance. The method for preserving the hand-drawn appearance of graphs described here is based on user observations and takes into account the context of the edge. The effectiveness of this algorithm was validated with a user study which suggests that people cannot distinguish the generated edges from hand-drawn edges.

1. Introduction

Using hand-drawn diagrams is a proven, easy and helpful technique in the early stages of design [6][12]. The process of sketching aids the communication of ideas, analysis of design, and creativity, while allowing alternative designs to be easily explored without concern for the cost of changes. Sketches are often used to create initial diagrammatic models of objects and processes: for example, UML diagrams, circuit diagrams and ER diagrams. Many of these diagrammatic forms are graphs that consist of a set of nodes with edges indicating the relationships between these nodes.

There are physical limitations to pencil and paper graph sketching. During design a diagram can often become convoluted and hard to understand as edges and nodes are added to the graph or are altered (Figure 1). To overcome this, the sketcher may need to go through a messy process of erasing and redrawing nodes and edges, or may need to restart the diagram altogether, as it is not possible to drag nodes around on paper.

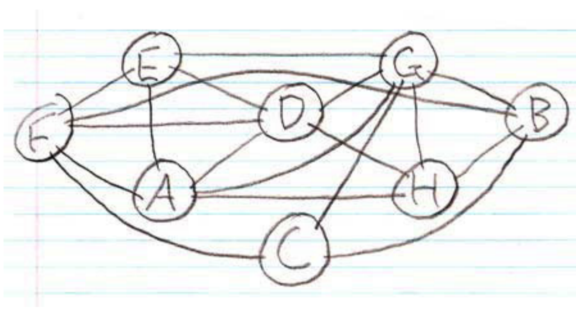


Figure 1: A hand drawn graph can quickly become messy without the ability to reposition nodes and edges

Computer-based sketch editing tools which allow the sketcher to reposition nodes manually by dragging them or which can apply an automatic layout algorithm can assist with the problem of tidying messy hand-drawn graphs.

Repositioning nodes in a sketched diagram introduces a new problem. How should an edge connected to a repositioned node appear after the node has been repositioned? This is the problem of *edge reflow.* Edge reflow is a common problem when editing a graph (e.g. [5]); here we consider the specific problem of edge reflow in when the graph is drawn using a sketch tool.

There are three particular challenges that need to be addressed when implementing sketch edge reflow:

(1) *The hand-drawn appearance of the diagram.* Hand-drawing is a simple and intuitive design process that is useful for brainstorming or communication. Crucial to the process of design is the hand-drawn appearance of a sketch. Sketch-editing graph tools should maintain the hand-drawn appearance of the graph, so as to best represent the creative ideas of the designer.

Designers' performance changes when working with designs of different visual fidelity [14]. Hand-drawn sketches permit designers to place emphasis on some areas while leaving others hazy and ambiguous; this helps exploration of alternative design ideas [6][12]. Bailey and Konstan compared a stylus based system against pencil and paper sketching and against Authorware [4]. They suggest that designers find hand-drawn designs most useful in the process of creation and that pencil and paper is most effective for exploring and communicating designs.

A sketched edge, once reflowed, should therefore leave the sketch with a natural hand-drawn appearance so that the advantages of sketching remain [14].

(2) *Intelligent reflow.* An edge does not exist in isolation: it is part of a graph of other nodes and edges. While previous implementations of graph sketch tools have used several techniques to retain the hand-drawn appearance of a reflowed edge (as described in Section 2 below), none have considered its context and its interaction with other visible elements of the diagram. For example, a reflowed edge in a graph should be adjusted so that does not pass through other nodes in the graph (Figure 2).

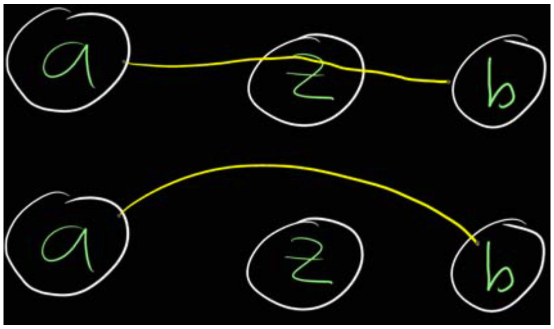


Figure 2: Intelligent reflow: The edge between a and b is best reflowed around z rather than going through it.

Such intelligent reflow is particularly important if automatic layout algorithms are to be applied to a hand-drawn diagram.

(3) *User-centered design and validation.* A reflow algorithm could be designed based on the intuitions of its designer. However, defining a reflow algorithm that does not take into account current sketching practices may result in a diagram that looks awkward or unnatural, or which follows conventions not typically used by sketchers.

A reflow algorithm is best designed after the graph drawing practices of sketchers have been observed. The results of this algorithm should be validated empirically so as to confirm that the reflowed edges are indistinguishable from hand-drawn edges, this proving its effectiveness.

In this paper, we present a new algorithm for determining the look of an edge after either its source or destination node has been repositioned in a graph sketch-editing tool. The advantages of this algorithm are:

- Its design was based on observations of users creating graphs;
- It is simple, being based on a library of hand-drawn edges;
- It preserves the hand-drawn appearance of an edge;
- It takes into account the context of the edge within the structure of the graph as a whole;
- Its output has been compared with hand-drawn edges, and its effectiveness validated with user studies.

Edge reflow is used in both the manual repositioning of a connected node and in the application of an automatic layout algorithm to the whole graph: the algorithm presented here can be used in both contexts.

2. Related Work

Various techniques have been explored to maintain the appearance of lines. Igarashi et al. [7] describe an 'As-rigid-as-possible' curve editing approach to hand-drawn line editing for cartoon characters. This process involved stretching a curve, by either scaling it to its new width or by a process of scale adjustment the curve acts as though it is a rigid object being pulled outward. These techniques are intuitive ways to morph an inflated graphic, but are not intended to preserve the hand-drawn appearance and may cause the curve to react in an unnatural way.

Arvo and Novins [2] explored edge reflow and preservation of sketch appearance within their blackboard style graphing system. They suggested techniques of edge reflow to preserve the hand-drawn appearance while also having the ink reflow in a predictable and intuitive manner. Their approach varied depending on whether: the new baseline (Figure 3) is shorter than the original baseline (compression); the new baseline is shorter than the original stroke length but is larger than the original baseline (stretch) the new baseline is longer than the original stroke length (over stretch).

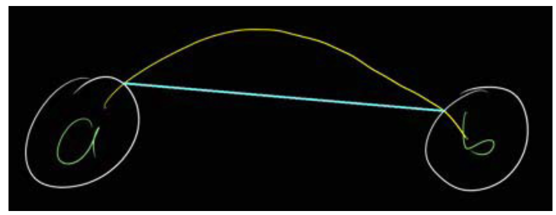


Figure 3: The straight line is the baseline of the arching connector stroke.

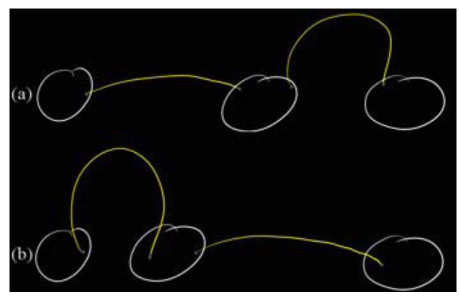


Figure 4: An illustration of the compression and stretch method [2]. When the centre node in drawing (a) is moved to the left, drawing (b) is created, with one connection stretched and interpolated with a straight line, the other compressed and interpolated with an ellipse.

Their metaphor is a piece of string. To compress the stroke points are interpolated with an ellipse (Figure 4a). To stretch a stroke (Figure 4b) it is interpolated with a straight line of the same length. The ratio of original stroke to straight line used for interpolation is decided by the amount of stretch: the greater the stretch the less the original stroke and the more of the straight line. This provides a smooth stretch and maintains the stroke length as though the stroke is a string being pulled and all the curves are slowly pulled out. When the total length of the stroke is less than the length of the straight line (over-stretched) the edge becomes perfectly straight. Arvo and Novins also presumed that it was important to maintain the relative position of the start point of an edge to the midpoint and boundary of the node and perform complex calculations to preserve this.

Reid et al [10] took a simpler approach to reflow maintaining more of the original appearance but avoiding straight edges. The stroke is rotated so its baseline lies flat, then scaled to the appropriate length along the horizontal and the height halved if it is above a given threshold. The problem with this approach is that strokes when heavily compressed look unnatural and stressed (Figure 5).

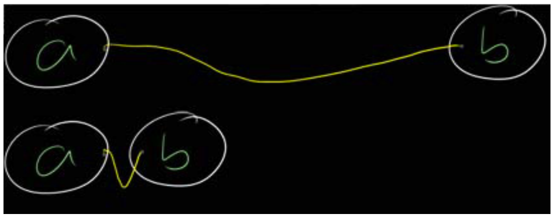


Figure 5: The top segment is the original; when node b is brought close to a, an unnatural looking edge is created

Ao et al. [3] considered appearance preservation in their network structure graph diagram sketching tool. When a container (node) is translated the associated connectors (edges) are moved with it. Two methods are reported: an uncomplicated scale to the new dimensions similar to [10]; and "stretching" by morphing of the stroke in a similar manner to [2]. This approach works in the majority of situations but in cases of extreme angles (elbows) it can cause loss of shape, in which case the scaling performs better.

All the above approaches are limited if we want to maintain the hand-drawn appearance in all situations: Arvo and Novins' [2] can result in perfectly straight edges while unnaturally curved edges occur in some cases with the Reid et al. [10] approach. Ao et al. [3] algorithms can result in both over-curved and over-straightened edges. Furthermore, none of these algorithms consider the context of the edge. This can result in edges cutting though nodes and intersecting with other edges; this compromises the hand-drawn appearance of the diagram.

3. Our edge reflow algorithm

3.1 Observational study

Before designing new algorithmic approaches to edge reflow we undertook a small informal observational study of how people create hand drawn graphs. Eight participants were asked to create two graphs each from a description and to then "tidy" the graph by repositioning the nodes.

Our observations of this process revealed that:

- When connecting two nodes, participants usually drew an approximately straight edge lying on the virtual line between the mid-points of the nodes, except when:
- (i) a third node lay on this virtual line (as between nodes p and o in Figure 6): in this case, the edge was routed around this node;
- (ii) there were two edges between a pair of nodes (as between nodes m and f in Figure 6): in this case, the edges were separated either side of this virtual line.
- Few participants could draw perfectly straight edges.
- The attachment points for an edge were close to the boundaries of the end nodes.
- People often crossed edges during construction but would eliminate crossings when tidying the graph. There were no incidences of edges crossing a node.

We did not observe any individual variation in edge style or the placement of node attachment points, and our observations and informal discussions with the participants suggest that they focus on the logical relationships between the nodes rather than the visual appearance of the edges.

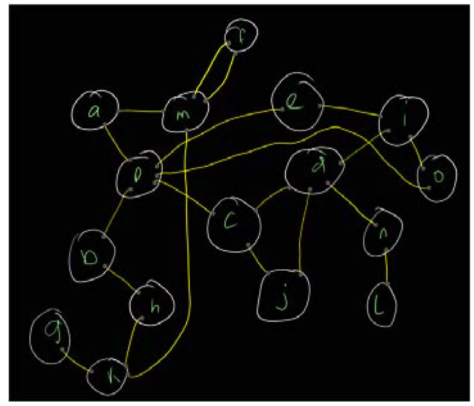


Figure 6: A typical initial construction of a graph

These observations formed the basis of the design of our new edge reflow algorithm to be used when nodes are repositioned either manually or automatically. Having observed no obvious user edge-drawing variations, the algorithm is designed to be generic, rather than specific to users' drawing styles.

3.2 Algorithm Design

Based on our observations, we suggest that, in order to maintain the hand-drawn appearance of the graph drawing, a reflowed edge should

- be a more-or-less straight edge (but not an exactly straight edge);
- lie approximately on the straight line between the centre points of the nodes.

This should be the case unless:

- there are multiple edges between a pair of nodes (in which case the edges should repel each other slightly), or
- the straight line between the nodes intersects with other nodes (in which case the edge should flow smoothly around the other nodes).

Our proposed solution is simple: when a node is repositioned, each of its connected edges is replaced by a generated edge.

The attachment points of these new edges are positioned at a small random offset from the point where the virtual straight line between two node centers crosses the node boundary, and at a small random offset within or outside the edge boundary.

Where there are multiple edges between a pair of nodes the new edges are repelled either side of the straight line and from each other so that there is visual separation.

Where the virtual straight line between an edge's nodes crosses other nodes, the generated edge is *flowed* around the node. Simply repelling the edge from the centre of the node results in an unnaturally curved edge with a flawless curve path. We explored a number of different functions to smooth the flow of the edge around a node, including a Gauss function [1], and a cosine function. We chose the cosine function (which we call *Context Reflow*), as it seems to give the most natural appearance, being not too sharp, and allowing multiple areas of force repulsion to be used without the stroke appearing too warped or disfigured. Figure 7 shows an example of each of these reflow approaches.

3.3 Implementation

There are four steps to reflowing an edge: the end points are established, a 'straightish' edge of appropriate length is generated, other node intersections are detected, and repulsion is used to push the edge away from other nodes.

The endpoints are established by identifiying the points of intersection between a straight line between the node centre points and the node boundaries (Figure 8, a and b) and then offsetting each by a small random amount.

The edge is generated from a small library of hand-drawn 'straightish' strokes of different lengths: this library was created and has been extended by many different users over the past two years. The new unique edge is generated from this library by random morphing between random pairs. The new edge is placed on the graph (Figure 8, edge ab).

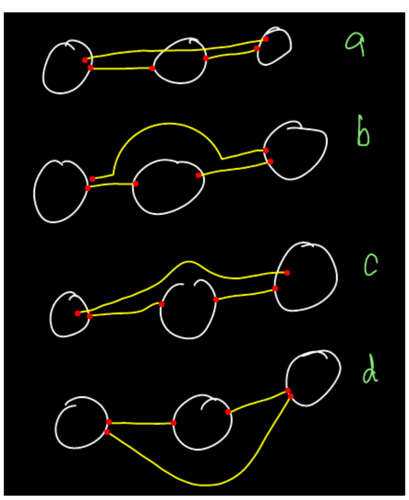


Figure 7: Node avoidance: a) no reflow, b) simple force, c) Gauss function force, d) Cosine function force (*Context Reflow).*

The new edge is then checked to see if it intersects with any other nodes. If there are no intersections the edge creation is complete. If intersections are detected the context reflow is applied.

The first step in the context reflow is to establish the exclusion zone. To do this all nodes that the edge intersects with are grouped and their bounding box found. The maximum distance for the displacement of the edge is calculated as the distance from the point on the edge nearest the midpoint of this exclusion zone to the most distant edge of the bounding box plus a padding value (Figure 8, c).

The maximum displacement is applied to this point. All other points are moved in relation to this point using formula (1). This allows the reflow to bend the edge over all the points in one arc. On a crowded graph the reflowed edge may cut through the corners of the bounding box; these cases are rare and cause few problems. A more sophisticated reflow would have to consider the individual bounding boxes of all the interfering nodes.

$$displacement \cdot Cos\left(\frac{\pi \cdot dist}{2\, max\, Dist}\right) \quad (1)$$

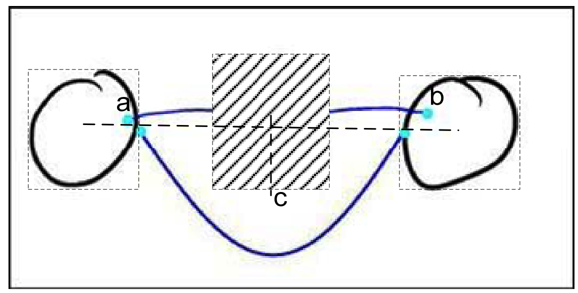


Figure 8: Context reflow: the edge is first placed between the connection points a b, and then reflowed around the bounding box of the intersecting nodes.

Where there are two edges between a pair of nodes the endpoints are pushed away from each other and the edges recreated between the new endpoints. This results in two 'straightish', and separated, edges between the nodes. Future enhancements to the system will include curving the edges away from each other.

4. Evaluation

Our aim was to produce reflowed edges that maintain the hand-drawn appearance of the graph drawing. Success can be determined by seeing whether the generated edges are indistiguishable from hand-drawn edges.

12 drawings of the same graph containing 10 nodes and 15 edges (referred to as D1-D12) were created by one of the authors. All the edges in D1 were hand-drawn (Figure 9). All the edges in D12 were system-generated, (Figure 10). The remaining D2-D10 drawings each had between 6 and 9 generated edges, with the remaining edges being hand-drawn (Figure 11).[1]

[1] A minor oversight meant that three of the drawings (D5, D6 and D7) were each missing one edge. This does not affect the validity of our experiment as these three edges are a small proportion of the total number of edges that each participant made judgements on.

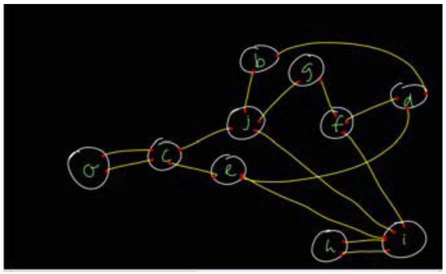


Figure 9: D1: All the edges are hand-drawn

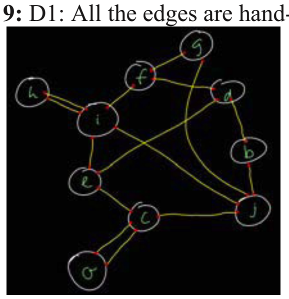


Figure 10: D12: All the edges are system-generated

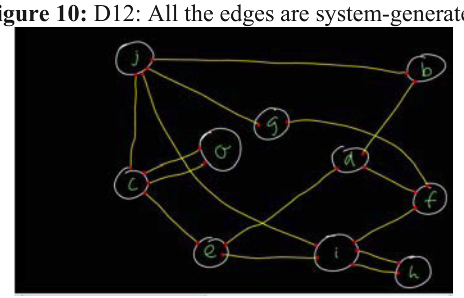


Figure 11: D8: Six edges are hand-drawn; nine edges are system-generated.

The following constraints were taken into account when creating these graph drawings:

- It is well-known that edge crossings can negatively affect perticipants' view and use of a graph drawing [8][9]. We did not want participants to be biased towards choosing edges that cross as the generated edges simply because they appeared awkward or anomalous. We therefore chose a non-planar graph which can be drawn with a minimum of two crossings. All 12 graph drawings had not more then four edge crossings: some of these were hand-drawn while some were generated.
- The nodes were all labelled, so that we could keep a record of which edges were system-generated. The same labels were used on all graph drawings, but different edges were chosen to be the ones that were system-generated.
- Some of the hand-drawn edges were deliberately drawn as curves (Figure 12).

Ten participants were shown all 12 graph drawings and asked to distinguish between hand-drawn and system-generated edges. They marked on the drawing with a pen, indicating those edges that they thought were system-generated with a C and those that they thought hand-drawn with a P. There was no time limit, and participants were encouraged to mark all the edges on all graph drawings.

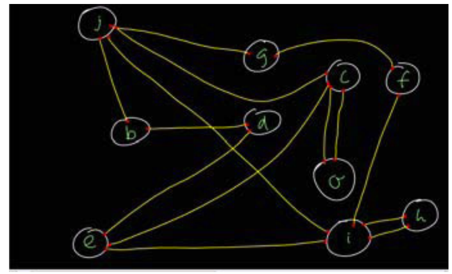


Figure 12: D5: The edge between nodes g and f has been hand-drawn.

The data collected was the number of correctly identified hand-drawn edges (HD), the number of correctly identified system-generated edges (SG), the number of hand-drawn edges incorrectly identified as system-generated edges (HDSG) and the number of system-generated edges incorrectly identified as hand-drawn (SGHD). Any informal comments made by the participants were also noted.

5. Results

Overall, 50.5% of the hand-drawn edges were classified as system-generated, and 52.6% of the system-generated edges were classified hand drawn.[2]

The data for the 'mixed-edge' drawings (D2-D11) was aggregated; the data for the 'control' drawings (D1, D12) was analysed separately (Table 1).

		D2-D11	D1	D12
HD (correct)	Mean	52.29	60.00	*NA*
	Max	75.00	86.67	*NA*
	Min	29.31	26.67	*NA*
HDSG (incorrect)	Mean	51.01	40.00	*NA*
	Max	75.41	73.33	*NA*
	Min	29.17	13.33	*NA*
SG (correct)	Mean	46.71	*NA*	49.63
	Max	59.72	*NA*	66.67
	Min	34.48	*NA*	26.67
SGHD (incorrect)	Mean	53.04	*NA*	50.37
	Max	65.52	*NA*	73.33
	Min	38.46	*NA*	33.33

Table 1: Mean, max and min percent of correct and incorrect classifications over all participants.

Paired t-tests were used to determine whether there was any significant difference between (a) the percentage of hand-drawn edges correctly identified (HD) and those incorrectly identified (HDSG) and (b) the percentage of system-generated edges correctly identified (SG) and those incorrectly identified (SGHD). In both cases, the probability of the classifications having been made simply by chance was high, and over the traditional p-value of 0.05 used for testing statistical significance (HD/HDSG: $p=0.452$; SG/SGHD: $p=0.117$). This proves that

[2] Although all participants were encouraged to classify all the edges, 10 of the 120 marked-up graphs had at least one missing label: these drawings were removed from the analysis.

participants could not distinguish between system-generated and hand-drawn edges, and that their decisions were as good as if they had been random.

We note, however, that the difference in the averages for the HD/HDSG comparison (1.28%) is less than that for SG/SGHD comparison (6.33%) by a factor of almost 5, suggesting that it was easier to correctly classify system-generated edges than hand-drawn ones. This observation is also shown in the narrower range between maximum and minimum percentages for identification of system generated edges than hand-drawn ones.

However, the two control drawings show that there was greater success with a completely hand-drawn diagram (mean 60.00%) than with a drawing with all edges system-generated (mean 49.63%).

The qualitative questionnaire data revealed that the following features were considered important for classification:

- The 'curviness' of the edge. Many participants said that curved edges were hand-drawn and straight edges system-generated.
- The 'kinkyness' of the edges. Smooth edges were considered as computer-generated, while those with 'bumps and curves' were classified as hand-drawn.
- The edge connectors. Edges with connectors close to the node boundary were labelled system-generated as 'the human dots are less accurate'.

General comments made it clear that the participants found this a very difficult task; e.g. "Nigh on impossible to make a decision. They all look the same", "For the majority of the lines, I just guessed."

6. Discussion

Our goal in this project is to retain the hand-drawn appearance of edges in a graph drawing sketching system when the manual or automatic repositioning of nodes requires that they be redrawn.

Our implementation of a *Context Reflow* algorithm is based on our observations of how people draw graphs. It is both simple and flexible, and in this paper we have demonstrated that its results are indistinguishable from hand-drawn edges.

Our results indicate further improvements to the edge-reflow algorithm; in particular, the placement of the edge connectors appears to be too precise to be comparable to hand-drawn edges.

We also need to consider how to adapt this algorithm to deal with several interfering nodes, rather than one. In addition, we are keen on investigating whether using a library of hand-drawn curved edges which may be morphed and adjusted at real time may be useful in edge reflow.

We have implemented other reflow algorithms and chose the *Context Reflow* one for this first empirical study based on our own intuition. An empirical comparison with the Gaussian method (Figure 7 (c)) and with enhanced versions of our *Context Reflow* method is needed in order to determine the best reflow algorithm for maintaining the hand-drawn appearance of the graph.

7. Conclusion

Many algorithms for edge reflow have been developed: the strength of the one reported here is that its design is based on observations of humans creating graphs, and that its success has been empirically validated.

Sketch design tools that do not preserve hand-drawn appearance do not fully utilize the stylus and interfere with one of the known advantages of pencil and paper. Our reflow algorithm is a validated approach to ensuring that the hand-drawn appearance of a graph is maintained after modification.

References

[1] Anton, H and R. C. Busby, Contemporary Linear Algebra, John Wiley & Sons, Inc, 2003.

[2] Arvo, C and K. Novins, Appearance-preserving manipulation of hand-drawn graphs, Graphite, ACM, 2006, pp. 61-68.

[3] Ao, X., Wang, X,. Jiang, W. and G. Dai, Structuring and Manipulating Hand-Sketched Diagrams, Sketch Based Interfaces and Modeling, Eurographics, Riverside, CA, USA, 2007

[4] Bailey, B. P. and J. A. Konstan, Are Informal Tools Better? Comparing DEMAIS, Pencil and Paper, and Authorware for Early Multimedia Design, CHI 2003, ACM, Ft Lauderdale, 2003, pp. 313-320.

[5] Dwyer, T., Marriott, K., and Wybrow, M. Dunnart: A Constraint-Based Network Diagram Authoring Tool. In GD 2008, pp 420-431.

[6] Goldschmidt, G., The backtalk of self-generated sketches, in J. S. Gero and B. Tversky, eds., Visual and spatial reasoning in design, Key Centre, University of Sydney, Sydney, 1999, pp. 163-184.

[7] Igarashi, T., Moscovich T, .and J. F. Hughes, As-Rigid-As-Possible Shape Manipulation, International Conference on Computer Graphics and Interactive Techniques, ACM New York, NY, USA, Los Angeles, California, 2005, pp. 1134 - 1141.

[8] Purchase, H, Carrington, D. and J.-A. Allder, Empirical Evaluation of Aesthetics-based Graph Layout Empirical Software Engineering, 7, 2002, pp. 233-255.

[9] Purchase, H., Which aesthetic has the greatest effect on human understanding? , Graph Drawing, Springer Berlin / Heidelberg, 1997, pp. 248-261.

[10] Reid, P., Hallett-Hook, F., Plimmer, B. and H. Purchase, Applying Layout Algorithms to Hand-drawn Graphs, OzCHI 2007: Entertaining User Interfaces ACM, Adelaide, 2007, pp. 203-206.

[11] Soni, BK., Thompson, JF., and N. P. Weatherill, Computer-Aided Geometric Design, Handbook of Grid Generation, CRC Press LLC, 1999.

[12] Tversky, B., What does drawing reveal about thinking, Visual and spatial reasoning in design, Cambridge, Mass., 1999, pp. 75-81.

[13] Ware, C. Purchase, H., Colpoys, L., and M. McGill, Cognitive measurements of graph aesthetics, Information Visualization 2002, pp. 103 – 110.

[14] Yeung, L., B. Plimmer, Lobb, B. Elliffe, D., 2008. Effect of Fidelity in Diagram Presentation HCI 2008. D. England. Liverpool, BCS. 1: 35-45.

ReCCO: An Interactive Application for Sketching Web Comics

Ricardo Lopes Manuel J. Fonseca
INESC-ID/IST/TU Lisbon, Lisbon, Portugal
ricardo.lopes@gmail.com, mjf@inesc-id.pt

Tiago Cardoso Nelson Silva
inEvo, I & D, Lisbon, Portugal
{tiago.cardoso,nelson.silva}@inevo.pt

Abstract

Over the last years, casual and amateur artists have been using the emerging Internet technologies to easily produce, share and distribute their web comics. However, easiness and quickness are not enough. These web applications to create comics still show very simple and static interaction methods, restricting users' experience.

With this paper, we propose a solution for creating web comics which builds upon and extends the traditional principles of creating traditional comics. Our approach combines sketch-based interaction to draw comics with shape and command recognition, computer assisted drawing and retrieval of previous visual elements. Experimental evaluation demonstrates that our approach allows users to create more appealing and richer comics with higher flexibility and efficiency than with existent applications.

1. Introduction

In recent years, the Internet has grown its influence on people's lives. The creation of comics has also been strongly affected by this technology with the number of amateur and inexperienced authors significantly increasing. These amateur artists can now create comics online in a fast and simple way, instantaneously sharing it with a global audience of readers. However, quickness and easiness in the creative process are not enough. Current web applications for creating comics online typically use the composition, arrangement and annotation of images as the most frequent technique for creation. In this method, images are selected from a pre-populated small database that users have no control of. Thus, available web applications are strongly restricting users' creativity and interaction by not allowing users to explore their abilities to create complex and rich comic strips, similar to those on paper. Additionally, these solutions do not automate some of the needed, but time-consuming, tasks, such as, the repetition of visual elements over several comic strips or the creation of simple but very common shapes.

To solve these problems we propose a new approach for creating comics online, which extends the principles of traditional comics creation by using sketch-based techniques for creating and retrieving visual elements. Our approach combines rich interaction mechanisms based on free hand sketching with shape and command recognition and retrieval techniques. With our approach, we aim at enabling users to create visually complex comic strips in a more flexible and efficient way, while maintaining the easiness and quickness of the process.

The rest of this paper is organized as follows: Section 2 provides an overview of related work in comics creation and shape recognition systems. In Section 3 we briefly present our approach, while in Section 4 we detail our solution by describing the calligraphic comics rich editor and its recognition and assisted drawing mechanisms. Section 5 shows and discusses experimental results from tests with users. Finally we discuss our conclusions and future work.

2. Related Work

Since our solution builds upon the principles of traditional comics, we started by analyzing these type of techniques, already compiled in specialized books [7, 2]. We observed that the comics medium has a specific visual nature defined throughout the years by different artists. They have been using the same visual conventions to represent information, such as, sound, speech, movement, emotions and time, which constitutes the vocabulary of a universal language for comics creation. In these books authors define the narrative structure of comics as a deliberate sequence of juxtaposed images, intended to convey information.

Regarding comics in digital format, we studied Williams et. al. work [15], called the ComicKit, which allows users to create comics by combining multiple visual elements available into a database. Users can pick the next element to insert into the story by searching images in the database. ComicKit uses the most frequent technique exploited in this type of systems, the combination and manipulation of predefined images. Comeks [11], Strip Generator [13] and Strip Creator [14] are other examples of this type of solution for creating comics online. However, these systems fail in allowing the creation of visually complex comics, since users' creativity is limited not by their skills, but by the set of available images and by system functionalities. Additionally, they fail in providing adequate support for the reuse of previous elements. Users can create new objects by combining multiple images, but they cannot reuse them in the same or in another comic strip.

Other solutions, like the Comics Sketch [5], try to stimulate users to create richer comic strips by supporting free hand drawing. Results show that users explore this kind of freedom and are able to draw more complex comics than with the previously described solutions. However, Comics Sketch has a very reduced set of editing operations, working as a simple digital sketch-pad, failing to help users on the creation of comics in a flexible and efficient way.

Some studied systems are taking the process of creating comics a bit further by giving suggestions to users. ComicKit suggests the next element to be picked, when the author selects an image to use. This system uses similarity and predefined subject-object relationships as the criteria for suggesting elements. When the user is not pleased with the suggestions made ComicKit allows a correction through a keyword-based retrieval mechanism. Results show that users found suggestions as an intelligent and helpful method to create comics. Also in the area of calligraphic interfaces are solutions that use a suggestion list to present possible elements for users to include in their drawings or to solve ambiguities during the drawing creation, such as Igarashi's Chateau [4] and Pereira's GIDeS [9]. Suggestion list is a kind of non-intrusive context-based dynamic menus containing the different interpretations of the sketched input. Sketch recognition systems, which sometimes are used in combination with a suggestion list, try to identify shapes and commands from users' strokes by analyzing features associated to them [3] or by analyzing how the stroke was drawn [10, 12, 8].

Looking at the majority of comics creation systems we can observe three things. First, these solutions are restricting user interaction and, as a consequence, the visual richness of resulting comics strips. Second, few systems offer computer assisted techniques, such as suggestion mechanisms or recognition techniques to increase the efficiency and flexibility of the creative process. Finally, there is no support for the reuse of visual elements, a need strongly inherent to the traditional comics creation process.

3. System Overview

Our approach offers a new methodology for creating comics online by combining free hand drawing and editing with sketch recognition and computer assisted drawing (CADr) techniques and a sketch-based retrieval mechanism based on calligraphic tags. We defined an architecture for a web application, which has two main components: the editor module, which is on the client side, and the retrieval module, which is a server application.

The editor module is responsible for the interaction with the user and allows the creation and edition of comics. It has a component-oriented architecture where each component can be viewed as a tool set with specific responsibilities and interface. The editor module is, at its core, a library of individual components that can be used as building blocks of a client web application. The combination of all these tools leads to our solution. Here, we stress two of these components: the Canvas and the Expectation List. The Canvas is the main sketching area and supports not only the creation of new comics, but also works as the interface with the retrieval and CADr mechanisms. The Expectation List is the component that presents possible visual elements produced by these two mechanisms. This component has a very important role when we consider the case of the CADr mechanism. The List not only shows results, but implements and supports all the logic behind this specific solution.

The retriever module is responsible for supporting the execution of all the mechanism and algorithms to retrieve previous drawings created by the user. Users interact with the retriever module through calligraphic queries that are matched against drawings in a database. Next, the most similar results are presented in the Expectation List.

4. Calligraphic Editor

To support the cornerstone of our solution, the rich creation and edition of comics, we created an editor module responsible for providing a large set of interaction techniques and operations. This combination leads to a new method for creating comics with more flexibility, efficiency and freedom than current web applications for comics creation.

4.1. Editor Overview

The editor module supports a set of requirements identified as capable of enabling users to create more complex comics in a simpler way. These requirements support free hand drawing, sketch transformations, visual layers, sketch grouping, textual elements, zoom, CADr and retrieval of previous elements.

The editor module has a component oriented architecture, which distributes these requirements by individual blocks. These combine user interface with functionality, and are independent and configurable. ReCCO, the web application for comics creation was built by instantiating and configuring all of these individual components. Due to the flexibility and extensibility of our component-oriented

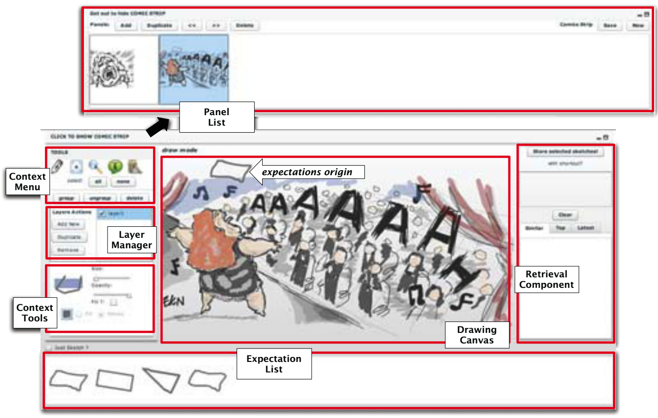


Figure 1. The different components of the Comics Editor.

module, we can build other applications with different requirements from our editor by instantiating a subset of the available blocks or even extending the editor with new ones. Figure 1 shows the various blocks of the editor interface. The Canvas is the main entity in our application and allows not only the drawing of new elements but also the seamless integration with the CADr and retrieval mechanisms.

In our recent work [6], we described our solution for the retrieval of previous elements in creating web comics. We proposed a new sketch-based retrieval scheme, where users can store and classify their drawings using calligraphic tags (or shortcuts). These calligraphic tags can then be used as queries to retrieve and reuse previously stored drawings, across different comic strips. The Canvas and the Expectation List (Figure 1) provide the interface between the editor and the retriever module, where the logic behind this scheme is implemented. The Canvas has this role because it is the component which captures users' mouse gestures and converts them either into final drawings or sketches which are candidates for retrieval or recognition. The Expectation List is the component responsible for observing the candidate sketches and executing the appropriate mechanisms. In the retrieval case this component communicates with the retriever module. In the recognition case, the Expectations List provides the interface and the mechanisms for CADr through sketch recognition. We will explain these techniques in the following sections.

The Context Menu is the component which allows users to control the application mode and decide, what to do with the sketches. This component is responsible for changing the active edition mode and accordant edition components, thus altering the behavior of the application. We have included edition modes for drawing, sketching commands, zooming, inserting/editing text and retrieving previous elements. The transformation of elements in its position, size and rotation is enabled in all edition modes.

The Panel List is the component which allows users to view and manage their comic strip. The Canvas and the Panel List communicate to update the current scene. For the internal representation (or Model) of created drawings and scenes we use the SVG format. This way, we maintain independence between the saved drawings and our web application. To achieve this goal, we defined our Canvas component as an SVG editor, with the ability do read, draw and edit SVG individual elements, in real time. This allows the communication between the Canvas, the Panel List or even the retriever module using exclusively this format. Indeed, with this solution, we can import any SVG element to the Canvas, use it directly or modify it to include in the comic strip. Finally, the Layer Manager allows the user to distribute drawings on a pile of overlapped visual layers.

4.2. Shape and Command Recognition

In our solution we included CADr mechanisms to assist users while drawing or transforming new content for the current comic frame. These mechanisms use geometric recognition algorithms to identify users' sketches and act accordingly. We use these algorithms to suggest new shapes and to support calligraphic-based commands in our application.

We adopted the CALI Library [3] for the ReCCO prototype to support shape and command recognition. CALI is a software library for calligraphic interfaces that provides support for the recognition of geometric shapes, such as, rectangles, circles, triangles, etc., independently of their size or rotation. The recognition algorithms use Fuzzy Logic and geometric features, combined with an extensible set of heuristics to classify sketches. To further extend the set of recognized shapes by CALI we adapted the Singularity [1] library to our solution. This library consists of a base set of classes for constructing single-segment and composite (piecewise) parametric curves. We converted CALI and Singularity libraries into Adobe Flex modules, since we developed our editor using this language.

Sketches drawn in the Canvas are processed and the correspondent geometric shapes are recognized. Identified shapes are then suggested as alternatives to replace the original sketch. The CALI library is able to identify circles, ellipses, lines, rectangles, triangles and closed polygons. The Singularity library identifies cubic bezier curves and closed curved shaped polygons. Figure 2 illustrates the combination of both recognizers in the identification of a sketch.

The CALI library is also able to recognize a special type of shapes, classified as gesture commands. When a scribble is recognized as a gesture command a specific Action is triggered and it performs a task on the Canvas and on the target

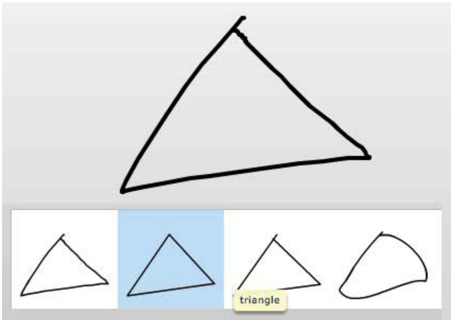


Figure 2. The sketch (above) is recognized as one of the (bellow) possible shapes: original, triangle, polyline and spline

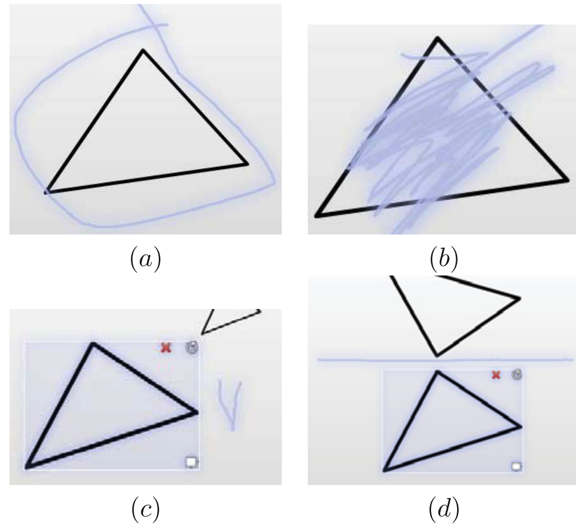


Figure 3. Recognized commands for: (a) select, (b) delete, (c) copy, (d) mirror copy.

elements. Our set of gestures and calligraphic actions allow the basic operations in the context of a drawing application. For instance, stroking over several drawings leads to deleting them; sketching a lasso around elements selects them; drawing a line creates the axis for performing a mirror copy of the selected drawings; and finally, sketching a 'V' character creates a copy of the selected drawing. Variations in the 'V' character size and rotation are recognized and have a direct relation to the size and rotation of the copy. Figure 3 illustrates these commands.

4.3. Computer Assisted Drawing

Our calligraphic editor, aimed at relieving users from some of the time and skill consuming tasks, while providing more natural and intelligent ways of interacting and creating. To accomplish these goals we adopted computer assisted drawing (CADr) techniques that provide usability gains to our solution. We implemented two mechanisms to support these goals: an intelligent suggestion list and calligraphic recognition (described above).

The suggestion list uses the shape recognition mechanism to identify and suggest a set of shapes that are geometrically similar to users' sketches. We aim to assist users in the task of drawing new elements and increasing the satisfaction with their own creations. Drawing perfect geometric shapes is no longer a time and patience consuming task for amateur users. Although we are suggesting "perfect" shapes to users we would point out that this mechanism is not restricting users' creativity, since these shapes are optional and are only suggested for simple geometric elements (e.g. circles, triangles, etc.). Furthermore, if users decide to adopt a suggested shape, they still can transform it as they do on a normal sketch.

Calligraphic actions are supported by the gesture commands recognition mechanism and assists users in interacting with drawings in the comic strip. Calligraphic actions can perform complex and flexible operations on specific drawings through the same interaction paradigm that users use while sketching them.

The Expectation List centralizes all the implementation of the calligraphic recognition and suggestion mechanisms. It listens to the sketching events in the canvas and responds to them. This behavior means that all of the CADr functionalities are executed exclusively on the client-side of our solution (editor module) allowing us to achieve good performance gains. Figure 4 shows the architecture of the Expectations List, regarding the CADr mechanisms.

The Expectation List has an Agents layer, responsible for processing and recognizing sketches created in the Canvas. The Expectation List invokes a specific agent, depending on the user's intentions or the application's context mode. If users sketch a shape, the CALI Agent and the Filter Agent (Singularity) are invoked by the Expectation List to identify

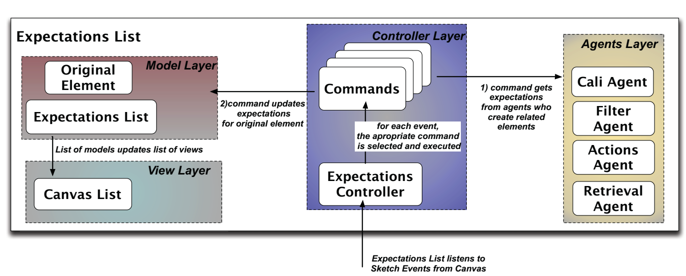


Figure 4. Expectations List architecture

the correct shape. To avoid ambiguity and provide better control suggestions are activated in the drawing mode and commands can only be performed in the actions mode. The Expectation List is responsible for compiling suggestions calculated by these two agents, present the results, capture the user's selection and, if necessary, replace the original sketch by the selected shape.

The Actions agent does not update the Expectation List since the result of the recognition process is a command and not a shape. After identifying the sketched command the operation is performed on the target element in the Canvas.

5. Experimental Evaluation

We performed experimental tests to compare our prototype, ReCCO, with an existing solution, Comics Sketch [5]. We intended to check if our editor and its CADr solutions improve the flexibility and efficiency of creating comics online. To that end, we asked participants to perform a task on both applications. The task consisted in the illustration of a predefined scripted story composed by six frames. The plot was created to encourage users to perform the task using all the components and mechanisms present in our solution.

Users created the same story in both applications. However, while half of them created first in ReCCO and then in the Comics Sketch (group R/CS), the other half did it in the reverse order (group CS/R). With this we tried to avoid biasing results, since users learn how to do the task from one application to the other. After users perform the task we did a questionnaire and an interview to collect their feedback about the applications. We selected 8 participants who represent the adequate user profile of our application. We preferred to perform tests with few "specialized" users, which are used and motivated to create comics (in an amateur environment), than to collect a larger number of users without any experience or motivation on comics creation.

Experimental results (see Figure 5) show that on average users spent 15 minutes less creating the complete comic strip in ReCCO than in Comics Sketch. From observation and interviews we notice that this difference is due to the edition, CADr and retrieval techniques present in our editor, since these allow a more powerful interaction.

We collected measures about the users' interaction with the different edition components and CADr techniques. However, we first defined a set of expected results through pre-conditioned pilot-tests, to get meaningful evaluation objectives. Figure 6 shows the results achieved for the edition components and mechanisms. From these results, we can observe that our edition and CADr components were explored by users as we expected. The average number of occurrences obtained for the edition and CADr operations were within the expected results. The number of transformations and deletions of elements were even significantly

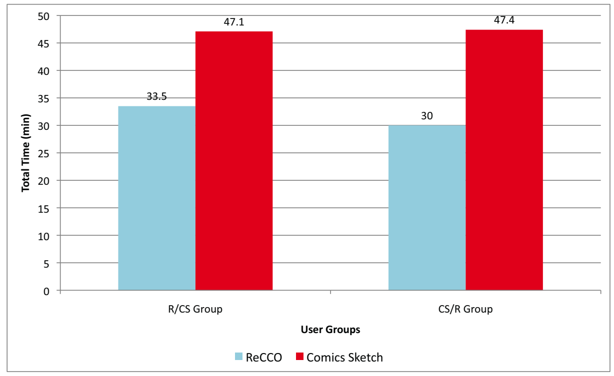


Figure 5. Creation times for the overall strip.

higher than expected showing users receptivity to a flexible rich edition interaction. However, the average of layers per each comic frame was lower than expected and close to 1. Maybe, the simplicity of the scripted story in this evaluation, did not encourage a high number of layers. Further tests would have to be conducted on this component.

We also measured users satisfaction with our prototype through questionnaires and interviews. The CADr mechanisms were the most commentated by users, who single them out (and the retrieval mechanisms) for their efficiency, easiness and innovation. In what regards usability, all the participants chose our prototype as their favorite. The main reasons for that choice were the acknowledgment of an easier and quicker creation process, supported by the retrieval and the CADr mechanisms. 75% of the participants considered the suggestion mechanism as essential and 87.5% think that calligraphic actions are an easier way of interaction.

Finally, we evaluated the creativity and richness of each comic strip from each application. Since creativity is a strong subjective concept, we chose to analyze this aspect

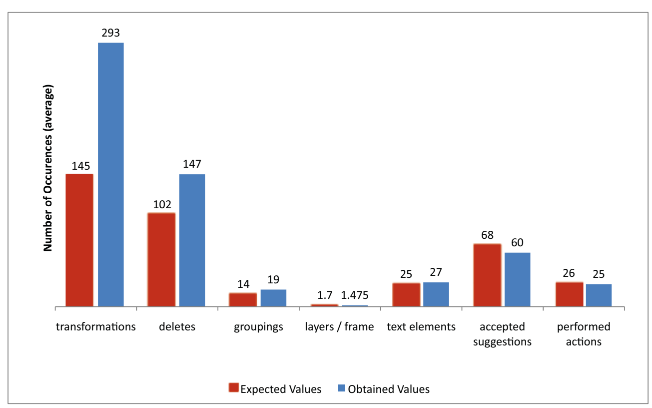


Figure 6. Evaluation of the edition methods.

Figure 7. Example of frames created with ReCCO (left) and Comics sketch (right).

on an individual basis and focusing on the personal opinion of each user about their own creations. The majority of the participants chose the ReCCO comic strip as their favorite and the most visually complex. Through observation we also noted that, with our prototype and due to the presence of more editing operations, users became more demanding and detail oriented with their comic strips, thus stimulating their creativity. Figure 7 shows the same frame created by the same user, using ReCCO and Comics Sketch.

6. Conclusions and Future Work

In this paper, we present a new approach for creating web comics. This solution combines a calligraphic rich editor based in free sketching with CADr techniques and a mechanism for the retrieval of previous elements. For the CADr techniques we use shapes and commands recognition algorithms as a way of assisting users to create and edit comics. Calligraphic commands assist users in the interaction with the system and a shape suggestion list aids on the task of drawing specific simple geometric shapes. Experimental results show that our approach enables users to create comics in a shorter time than with current solutions, while achieving visually rich and complex scenes.

The rich edition (with calligraphic actions) and the suggestion mechanisms were fully explored, with users quickly adopting this kind of interaction. We concluded that this kind of assistance to the creative process did not restrict users from creating visually rich content. In fact, the combination of these edition and suggestion mechanisms enables users to create comics while being simultaneously faster and more (visually) ambitious. We can claim that our approach liberates the user to create better comics in a more flexible, efficient and satisfying way.

In what regards future work, several directions were identified for the edition and CADr techniques, mainly in the actions or suggestion mechanisms. Considering the actions usability, we can improve this mechanism with more commands (copy without variations, undo, etc) and facilitate its integration with other edition context modes. Regarding suggestions, we think that the set of recognized shapes could be augmented with specific shapes from the traditional comics vocabulary (e.g. speech balloons). The repeated use of these shapes in a comics story encourages us to include them in a suggestion mechanism.

References

[1] J. Amstrong. Singularity - parametric curve library. *http://www.algorithmist.net/*, 2008. as seen in September 2008.
[2] W. Eisner. *Graphic Storytelling*. Poorhouse Press, 1996.
[3] M. J. Fonseca, C. Pimentel, and J. A. Jorge. CALI: An Online Scribble Recognizer for Calligraphic Interfaces. In *2002 AAAI Spring Symposium - Sketch Understanding*, pages 51–58, Palo Alto, USA, Mar. 2002.
[4] T. Igarashi and J. F. Hughes. A suggestive interface for 3d drawing. In *14th annual ACM symposium on User interface software and technology (UIST'01)*, pages 173–181. ACM, 2001.
[5] inEvo. Comics sketch. *http://mainada.net/comics*, 2008. as seen in December 2008.
[6] R. Lopes, T. Cardoso, N. Silva, and M. J. Fonseca. Calligraphic shortcuts for comics creation. In *9th International Symposium on Smart Graphics*, volume 5531 of *LNCS*, pages 223–232. Springer, 2009.
[7] S. Mccloud. *Understanding Comics*. Perennial Currents, 1994.
[8] B. Paulson and T. Hammond. Paleosketch: accurate primitive sketch recognition and beautification. In *IUI '08: 13th International conference on Intelligent user interfaces*, pages 1–10. ACM, 2008.
[9] J. P. Pereira, J. A. Jorge, V. A. Branco, and F. N. Ferreira. Calligraphic interfaces: Mixed metaphors for design. In *10th Int. Workshop on the Design, Specification and Verification of Interactive Systems (DSV-IS'03)*, June 2003.
[10] D. Rubine. Specifying gestures by example. *SIGGRAPH'91: 18th annual conference on Computer graphics and interactive techniques*, 25(4):329–337, 1991.
[11] A. Salovaara. Appropriation of a mms-based comic creator: from system functionalities to resources for action. In *CHI '07: SIGCHI conference on Human factors in computing systems*, pages 1117–1126. ACM, 2007.
[12] T. M. Sezgin, T. Stahovich, and R. Davis. Sketch based interfaces: early processing for sketch understanding. In *SIGGRAPH '07: ACM SIGGRAPH 2007 courses*, page 37. ACM, 2007.
[13] T. F. Studios. Strip generator. *http://stripgenerator.com*, 2007. as seen in October 2007.
[14] B. Turcotte. Strip creator. *http://www.stripcreator.com*, 2007. as seen in October 2007.
[15] R. Williams, B. Barry, and P. Singh. Comic kit: acquiring story scripts using common sense feedback. In *IUI '05: 10th international conference on Intelligent user interfaces*, pages 302–304. ACM, 2005.

Performances of Multiple-Selection Enabled Menus in Soft Keyboards

Gennaro Costagliola, Vittorio Fuccella
Department of Mathematics and Informatics
University of Salerno
{gencos,vfuccella}@unisa.it

Michele Di Capua, Giovanni Guardi
Unlimited Software
md@unlimitedsoftware.it, g.guardi@u-s.it

Abstract

In this paper we present a new text entry approach based on soft keyboards and a study on its expected performances. The basic idea of the approach is to enhance the keyboard with a menu whose items (characters) can be selected more than once through a single pointer (pen, finger, etc.) stroke, in order to enter particularly frequent text patterns. The gesture to enter these patterns is similar to that used for Cirrin. Furthermore, using well-established methods for measuring the time required for performing simple interactions with pointers on touchscreens, we have evaluated the expected performances of our method. The results, obtained through a text entry simulation, are encouraging with respect to those measured for previous approaches based on menu-augmented keyboard.

1. Introduction

Fast text entry in mobile devices is still a challenging issue. Soft keyboards on touchscreen are still largely employed on palmtop and smart phones. Soft keyboards enable text input through *tapping*. Each *tap* corresponds to a single character input. Even though more efficient keyboard layouts exist [10, 18], the QWERTY layout is still the most used, due to the users' familiarity acquired in the use of desktop and laptop computer keyboards. Other layouts do not meet a wide audience, probably due to the difficulty to become familiar with an unknown layout [11]. In order to accelerate text entry, menu augmented keyboards [5] have been introduced. They aim at improving text entry performances by enabling text entry through *flicks*, in addition to the classical *tap*. A *flick* corresponds to the input of a digraph in which the first letter is identified by the starting point of the interaction, the second by its direction.

In this paper we present a new text entry approach on menu-augmented soft keyboards and a theoretical study on its performances. The approach is based on the following idea: the menu items can be selected more than once, in order to enter particularly frequent text patterns with a single pointer stroke. The interaction to enter a text unit, by selecting a sequence of menu items, is similar to that used for *Cirrin* [12]. Users familiar with a given keyboard layout must not learn a new one: they just have to acquaint themselves with the new type of interaction.

Furthermore, the method has the advantage to significantly reduce the number of strokes (taps and gestures) necessary to enter text. Compared to past methods [5], which only enable input of digraphs, it enables the input of longer string chunks with a single stroke. It is worth noting that the user can choose whether showing the menu or not, eventually switching from *novice* to *expert mode*.

The paper also presents a theoretical analysis on the expected performances in terms of speed obtainable by an ideal user. The analysis model uses Fitts' Law [3] and Hick-Hyman Law [4] for evaluating performances for novice and expert users on the the QWERTY keyboard layout. Based on this model we built a text entry simulation and measured the expected performances of our approach showing that these are slightly better than those measured for previous approaches based on menu-augmented keyboard.

The rest of the paper is organized as follows: the next section contains a brief survey on text entry methods with soft keyboards, focusing mostly on the methods related to ours, and summarizes some basic concepts about performance analysis; section 3 gives an explanation of the proposed approach; in section 4 we present the analysis on the expected performances; lastly, some final remarks conclude the paper.

2. Related Work

2.1. Entry Methods

The spread of mobile computers and smart phones equipped with touch screens has attracted the interest of researchers on the problem of text entry in such devices. Several methods for accelerating the text entry task have been proposed. Some of the main directions followed by researchers are:

- proposal of *more efficient* keyboard layouts;
- proposal of interaction types minimizing the number of strokes for text entry.

In the former case, the research is mainly based on the idea that the keyboard layouts should minimize the distance between characters with a high probability of being consecutive in the words of target languages. To this aim, the frequency of digraphs in target languages (e.g. English) has been analyzed in texts and reported in tables [15]. The studies have resulted in the proposal of specific layouts, such as OPTI [10], Fitaly [16], Atomik [19] and Metropolis [18].

In the latter case, methods enabling text entry through interaction types different from *tapping* have been proposed. An example is the selection of pie menu items. Among these methods, the earliest, such as *T-Cube* [17], enabled input of single characters per menu item selection, through a *flick* interaction. Then, researchers noticed that enabling the menu item selection on a virtual keyboard it is possible to input digraphs through a single stroke. The first character is located on the keyboard and is identified by the starting point of the stroke. The second is located in the menu item and is identified by flick's direction.

Other methods enable the input of single characters with a single stroke without the use of a virtual keyboard (*unistroke* alphabets, such as *Graffiti* [2]), or enabling the input of entire words through a single stroke (word-level unistroke) on ad-hoc designed keyboard layouts. Among these, we can mention *Cirrin* [12] and *Quikwriting* [14]. In the above methods, the objective is pursued by arranging the keys in a special layout which facilitates the input of whole words without ambiguity. More recent methods enable the use of word-level unistroke still keeping keyboard layouts familiar to the user, as QWERTY. In those cases, e.g. in [6] and [20], it is necessary the application of sketch recognition and dictionary-based disambiguation methods.

We will now briefly describe *Cirrin*, since our method uses a similar interaction to enter text. *Cirrin* is an intuitive word-level unistroke text entry method on virtual keyboard proposed by Mankoff and Abowd in 1998. The keys are arranged in a circle, as shown in figure 1. In order to enter a word, the user, beginning from the middle of the circle, simply traces out a path that crosses the circumference at points corresponding to the characters of the word, in the right order. A space character is automatically inserted as soon as the pen is lifted. The arrangement of the letters in the circle is such to minimize the average distance the pen travels to write a word. *Cirrin* has shown itself to be about as fast as classical QWERTY virtual keyboards and is particularly suitable for mobile computers.

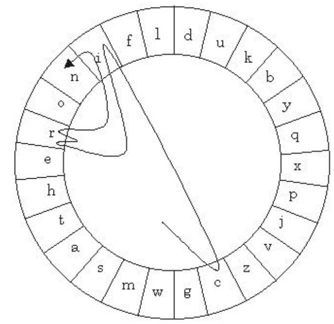


Figure 1. The Cirrin text entry method.

2.2. Performance Analysis

Performances of a text entry method can be evaluated from the points of view of speed and accuracy. Speed is measured in terms of *characters per second* (*cps*) or, more frequently, of *words per minute* (*wpm*). A word is defined as a sequence of five characters (regardless of whether they are letters, punctuation or spaces), thus we can convert *cps* in *wpm* by multiplying *cps* by 12. Accuracy can be measured as the percentage of errors in typing a phrase out of the number of characters of the phrase. For a more detailed insight on performance analysis, the reader should refer to [9].

While accuracy can be only measured in trials, speed can also be estimated with theoretical models. The models can be applied to measure both the novice and the expert performances. For the latter users, who are supposed to already know the keyboard layout, only movement efficiency must be considered, while for the former users, also the time needed for visual exploration to search for the destination key (or menu item) should be taken into account. To predict movement efficiency, Fitts' Law [3] is used. According to the Shannon formulation, [8], of the law, for movement along a single dimension, the average time T to complete the movement can be calculated as:

$$MT = a + b * log_2(D/W + 1) \tag{1}$$

where:

- a represents the start/stop time of the device and b stands for the inherent speed of the device. These constants can be determined experimentally by fitting a straight line to measured data.
- D is the distance from the starting point to the center of the target.
- W is the width of the target measured along the axis of motion.

To predict the Reaction Time (the time required to make a choice among n items), Hick-Hyman Law [4] is used. The law is expressed as follows:

$$RT = a' + b' * log_2(n) \tag{2}$$

where n is the number of items to choose from. The coefficients a' and b' are slope and intercept constants, similar to the coefficients a and b in Fitts' law.

3. The approach

The approach is an improvement of the classical stylus keyboarding. The improvement lays in enabling gesturing on the keyboard, besides the classical *tapping* interaction: the gesture is a drag of the pointer inside a menu containing a restricted set of frequent characters, one per menu item. The menu is shown as soon as the pointer is pressed on a key and disappears as soon as the user releases the pointer. The menu is displayed around the pressed key: menu items are arranged one per key side. More precisely, the internal perimeter of the menu coincides to the external perimeter of the key. For instance, if the character has a squared (or rectangular) shape, the menu will allow four different character choices.

While the menu is shown, it is possible to sweep out a gesture that touches only the desired characters in succession by dragging the pointer, without lifting it. This gesture is almost the same as for the *Cirrin* method. In *Cirrin* a space character is always automatically inserted at the end of the gesture, since the word is supposed to be completed with just one stroke. Since in our method text entry units are not always complete words, the space character can be inserted or not, according to the final position of the pointer when it is released: the adopted convention is to add a space at the end of the text unit only if the pointer is lifted after returning inside the character key area. Otherwise, if the pointer is lifted in the menu area, the string finishes with the last selected character. Note that in all the other cases the space bar is directly used.

Let us consider a soft keyboard augmented with a menu containing n characters $x_1, \ldots, x_n$. With our method, with a single stroke we can enter a text unit described by the following regular expression (for a short reference manual on regular expressions, see [13]):

$$.[x_1 x_2 \ldots x_n] + [\ \ ]? \tag{3}$$

The above pattern matches any text unit starting with any character (specified in (3) by the starting '.') chained to a sequence of one or more of the $x_1 x_2 \ldots x_n$ characters, (specified by $[x_1 x_2 \ldots x_n]$+) eventually ending with a space character, ([ ]?).

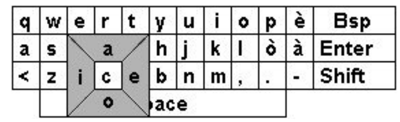


Figure 2. The QWERTY keyboard layout augmented with a menu.

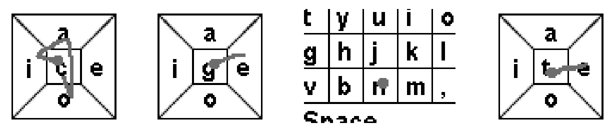


Figure 3. The 4 strokes needed to enter the text 'ciao gente'.

Our approach can be instantiated by associating each of the vowels 'a', 'e', 'i' and 'o' to one of the sides of the squared character keys. This choice is supported by the fact that these vowels are more easily remembered by users and that they are among the most frequent letters in many languages.

Arranging the menu items as shown in figure 2, the interaction sequence necessary to entry the Italian text *ciao gente* (*hello folks*, in English) is shown in figure 3. The string is ten characters long but it can be entered with a sequence of four strokes (taps or gestures). The strokes correspond to the input of the following sequence of text units {*ciao* }{*ge*}{*n*}{*te*}. Three text units out of four are entered through a gesture and only one through a *tap*. With the first stroke we can enter up to five characters.

Generalizing, the method allows to select a small number of frequent characters (4 with square keys, 6 with hexagonal keys and 8 with octagonal keys). The method can be instantiated differently for each specific language. In particular, the following parameters should be chosen appropriately:

- The number of menu items should be the result of a compromise: the highest this number, the highest the frequency of the match of the regular expression in (3) is. Conversely, the smaller this number, the faster the learning of the method by novice users is.

- The choice of the letters associated to menu items should take into account the frequency of the matches of the regular expression shown in (3), and users' learning preferences (vowels could be remembered more easily).

4. Expected performances

In this section we describe the theoretical analysis we have carried out on the expected performances in terms of speed obtainable by ideal novice (lower bound) and expert (upper bound) users. The users are supposed to be familiar with the keyboard layout. The performances have been measured for a user writing in Italian language. We expect similar results for other Romance languages, since they have similar letter and pattern frequency. The proposed evaluation model is an extension of the one used by Soukoreff and MacKenzie in [15] for evaluating virtual keyboard performances.

4.1 Reaction and movement time estimation

Since our approach enables two kinds of stroke (tapping and gesturing), our model separately evaluates the time needed for each of them. The former is the classical stroke used to enter a character located on the keyboard, while the latter is used to enter a text unit whose characters are located in the menu. The time required to perform each of them is calculated as follows:

Tapping. Since we assume that the users are already familiar with the keyboard, the required time is only given by the time needed to move the pointer from the previous key to the current one. We calculate the time to enter a c character located on the keyboard with the following formula:

$$Tk(c) = MT(c) \tag{4}$$

Where $MT(c)$ is the movement time to reach character c measured with Fitts' Law shown in (1).

Gesturing. In this case we consider both the novice and the expert performances. The time required for experts is only given by the time needed to drag the pointer among menu items, while for novices, we have to add the time necessary to visually scan the menu. This value is added every time a new menu item must be selected to enter the corresponding character, except for characters already selected in the same gesture. Summarizing, we calculate the time to enter a tu text unit using the menu with the following formula:

$$Tm(tu) = \begin{cases} MT(tu) & \text{if experts} \\ MT(tu) + RT(tu) & \text{if novices} \end{cases}$$

Where $MT(tu)$ is the movement time to enter the text unit tu, measured with Fitts' Law shown in (1); $RT(tu)$ is the reaction time (visual scan of the menu) to enter the text unit tu, measured with Hick-Hyman Law shown in (2).

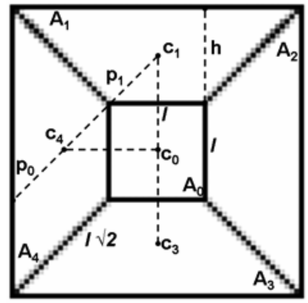


Figure 4. A model of menu with 4 items.

$MT(c)$ can be calculated with Fitts' Law. In particular, we use the following values for parameters:

- for intercept and slope, we use the values experimentally obtained by [7] for pointing tasks on tablet devices: $a = -55$; $b = 204$.
- for distance A, we use the distance between the centers of the previously pressed key and c;
- for W, the width of the target, we use the length of the side of the squared key.

It is worth noting that the measured time does not depend on the size of the keyboard, since both A and W grow proportionally with it.

$MT(tu)$ is calculated as the sum of the times to enter the single characters in tu. Supposing that tu is the concatenation of n characters $x_0 \ldots x_n$, then $MT(tu) = MT(x_0) + \ldots + MT(x_{n-1})$.

In particular, the movement time for the first character of the text unit, $MT(x_0)$, is calculated as done above for $MT(c)$, since it is located in the keyboard. The calculation of the time to select a character corresponding to a menu item $MT(x_i)$, with $1 \leq i \leq n-2$, i.e., is more complex and requires the analysis of the different movement types that can occur when dragging the pointer inside the menu.

Figure 4 shows a model of a key with a four items menu. We will refer to the internal area of the key with A_0, its center with c_0; the upper, eastern, lower and western areas are referred to with A_1, A_2, A_3 and A_4, their centers with c_1, c_2, c_3 and c_4. The key area is squared, while that of menu items has the shape of an isosceles trapezoid, whose smaller base coincides with the key side and whose base angles are 135 degrees. We also assume that the height of the trapezoid has the same length of the key side. The first movement, required to select the second letter of the text entry unit, is a drag from the center of the key to the center of a menu item, i.e. A_1. To estimate time, we apply Fitts' Law, considering that:

1. The measures of D and of the width of the target W are, respectively, the length of the segment c_0c_1 and of the height h of the A_0 trapezoid. They both have the same value l, as a consequence of simple geometrical considerations on the shapes shown in figure 4. This property does not hold for hexagonal and octagonal menu: the ratio D/W for them has, respectively, the value of 0.67 and 0.83.

2. for intercept and slope, we use the values experimentally obtained by [7] for dragging tasks on tablet devices: $a = -27$; $b = 276$.

Thus, by applying Fitts' Law, $MT(x_1) = a + b * lg2(2) = a + b$. To analyze the time for movements following the first one, $MT(x_i)$; $2 \leq i \leq n-2$, we have to analyze three possible cases:

a) movement between contiguous menu items;

b) movement between opposite menu items;

c) other movements (included the same menu item).

In case of type a) movements (i.e. between A_1 and A_4), D and W are, respectively, the length of the segment c_1c_4 and of the segment p_0p_1. They have the same length, thus by applying Fitts' Law, $MT(x_i) = a + b * log_2(2) = a + b$.

In case of type b) movements (i.e. between A_1 and A_3), D and W are, respectively, the length of the segment c_1c_3 and of the height of the A_3 trapezoid. D is twice W, thus by applying Fitts' Law, $MT(x_i) = a + b * log_2(3)$.

In case of type c) movements, the movement is regarded as composed of two simpler movements: a return in the key area followed by a new item selection, thus by applying Fitts' Law, $MT(x_i) = 2(a + b * log_2(2)) = 2(a + b)$.

Finally, the last movement can be a further character selection or a space. In the former case, $MT(x_{n-1})$ can be calculated as for the previous case. In the latter case, it is a return in the key area. It can be calculated by applying Fitts' Law, $MT(x_{n-1}) = a + b * log_2(2) = a + b$;

The time for menu scan can be calculated by applying the Hick-Hyman Law shown in (2). In particular, we use the following values for parameters:

- for intercept and slope, we use the values reported by [15]: $a' = 0$; $b' = 200$.

- for n, we use the number of menu items.

To calculate $RT(tu)$, we should add the above value for all the movements between menu items, except for movements for entering characters already selected in the same gesture. The above analysis, performed for movements inside a four items menu, is also valid for menus with more items.

4.2 Simulation

A software simulator has been developed, in order to calculate the time required to enter the text contained in an input text file. The simulator also accepts as input an XML file containing the definition of the keyboard layout. The output is a table which reports the time (in words per minutes) required to enter the text for each keyboard layout, both in its basic version and in the menu-augmented one.

Briefly, the simulator is equipped with a text scanner and a *TimeCalculator* module. The former scans the input text file and passes text units (single characters or strings matching using the regular expression (3)) to the latter. The *TimeCalculator* applies the rules described above in this section to calculate the time required to enter the text units.

4.3 Interpretation of results

The final result is a comparison of the performances on QWERTY layout. Menus with both single selection and multiple selection enabled with 4, 6, and 8 keys have been tested through the simulator. We have calculated the time required to enter the text contained in a corpus of Italian text, 58 books and journal articles published between years 1949 and 1996, freely downloadable from *Biblioteca della Scuola Normale* website [1]. The process of assigning the characters to the menu items has been aimed at minimizing movement times. In particular, from the calculations above, we know that type a) movements are faster than the other types. Thus, an optimal character arrangement should maximize movements between adjacent items.

Table 1 shows the results obtained by running the simulator with the text corpus on the QWERTY layout augmented with menu containing 4, 6 and 8 items. The second column in the table reports the sequence of the characters assigned to menu items, starting from the upper menu item and proceeding clockwise. The chosen letters are the most frequent in the Italian language. There is no warrant that this is the best choice for menu-augmented keyboards.

The arrangement has been performed trying to maximize movements between adjacent menu items. The following columns report the text entry speed in wpm: the third column reports the performances of the keyboard without any menu; the fourth column reports the performances of novice and expert users with a classical menu-augmented keyboard (one menu-item selection per time); lastly, the fifth column reports the performances of novice and expert users, with a multiple-selection enabled menu-augmented keyboard.

The results of our theoretical study are the following: for those users familiar with QWERTY layout, menu-augmented keyboards enable faster text entry speed when users become familiar with the menu layout too. Nevertheless, before acquainting themselves with it, their perfor-

Num of Items	char arrangement	QWERTY	Single Sel. Menu	Multiple Sel. Menu
4	aeoi	35.41	30.11 → 42.13	30.16 → 43.43
6	aenoli	35.41	26.87 → 41.58	25.21 → 43.30
8	aetnolri	35.41	24.78 → 40.97	20.35 → 38.94

Table 1. Performances (in words per minute) of menu augmented keyboards with QWERTY layout

mances are consistently lower. Our approach shows slightly better performances than classical menu-augmented keyboards with trained users with small menus (4 and 6 menu items). In particular, among our simulations, the fastest text entry has been obtained with the 4 items multiple-selection menu (43.32 wpm, with an improvement of 22.4% compared to simple QWERTY). Nevertheless, this result is not better than those obtained by other authors with optimized keyboard layout, such as *Fitaly* and *Opti II*.

Lastly, a result that could surprise the reader is that smaller menus have better performances than larger ones. We argue that this depends on the choice of the characters assigned to menu items. Other factors than character frequency can influence the performances of the method, such as the frequency of matches of the (3), and the keyboard layout: the character assigned to peripheral keyboard keys should be best candidates for the assignment to menu items, since their average distance from other keys is higher. Thus, their presence in the menu can help minimizing the pointer movements.

5. Conclusions and further research

We have presented a new text entry approach based on soft keyboards. The approach is an improvement of menu-augmented keyboards, consisting in the possibility of performing multiple selection of its items with a single pointer stroke. The approach has been theoretically evaluated in terms of text entry speed with an Italian text corpus. According to our simulation, expert users can input text faster than with classical QWERTY keyboard and also with QWERTY augmented with classical single selection menus, under certain circumstances. The above results encourage us to plan an experiment with human users, in order to measure the performances in real situations, from both the points of view of speed and accuracy. Before performing the experiment, further studies will be aimed at tuning our approach, by associating different characters to menu items and different menu instances to single keys.

References

[1] Biblioteca della 'scuola normale': Dizionari e corpora linguistici (in italian). http://biblio.sns.it/.

[2] C. H. Blickenstorfer. Graffiti: Wow!!!! *Pen Computing Magazine*, pages 30–31, 1995.

[3] P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. 1954. *J. of Exp. Psychology General*, 121(3):262–269, 1992.

[4] R. Hyman. Stimulus information as a determinant of reaction time. *J. of Exp. Psychology*, 45:188–196, 1953.

[5] P. Isokoski. Performance of menu-augmented soft keyboards. In *Proc. of CHI '04*, pages 423–430, New York, NY, USA, 2004. ACM.

[6] P.-O. Kristensson and S. Zhai. Shark2: a large vocabulary shorthand writing system for pen-based computers. In *Proc. of UIST '04*, pages 43–52, NY, USA, 2004. ACM.

[7] I. MacKenzie, A. Sellen, and W. Buxton. A comparison of input devices for elemental pointing and dragging tasks. In *Proc. of CHI 91*, pages 161–166. New York, 1991.

[8] I. S. MacKenzie. A note on the information-theoretic basis for fitts' law. *J. of Motor Behavior*, 21:323–330, 1989.

[9] I. S. MacKenzie and R. W. Soukoreff. Text entry for mobile computing: Models and methods, theory and practice. *Human-Computer Interaction*, 17:147–198., 2002.

[10] I. S. MacKenzie and S. X. Zhang. The design and evaluation of a high-performance soft keyboard. In *Proc. of CHI '99*, pages 25–31, New York, NY, USA, 1999. ACM.

[11] L. Magnien, J. L. Bouraoui, and N. Vigouroux. Mobile devices: soft keyboard text-entry enhanced by visual cues. In *Proc. of UbiMob '04*, pages 158–165, NY, USA, 2004. ACM.

[12] J. Mankoff and G. D. Abowd. Cirrin: a word-level unistroke keyboard for pen input. In *Proc. of the ACM UIST 98*, pages 213 – 214. ACM, 1998.

[13] The Open Group. *Regular Expressions*, the single unix specification, version 2 edition, 1997.

[14] K. Perlin. Quikwriting: continuous stylus-based text entry. In *Proc. of the ACM UIST 98*, pages 215 – 216. ACM, 1998.

[15] R. W. Soukoreff and I. S. MacKenzie. Theoretical upper and lower bounds on typing speed using a stylus and soft keyboard. *Behaviour and I.T.*, 14(6):370 – 379, 1995.

[16] Textware solutions: The fitaly one-finger keyboard. 1998. http://fitaly.com/fitaly/fitaly.htm.

[17] D. Venolia and F. Neiberg. T-cube: a fast, self-disclosing pen-based alphabet. In *Proc. of CHI '94*, pages 265–270, New York, NY, USA, 1994. ACM.

[18] S. Zhai, M. Hunter, and B. A. Smith. The metropolis keyboard - an exploration of quantitative techniques for virtual keyboard design. In *Proc. of UIST '00*, pages 119–128, New York, NY, USA, 2000. ACM.

[19] S. Zhai, M. Hunter, and B. A. Smith. Performance optimization of virtual keyboards. *Human-Computer Interaction*, 17:229–270, 2002.

[20] S. Zhai and P.-O. Kristensson. Shorthand writing on stylus keyboard. In *Proc. of CHI '03*, pages 97–104, New York, NY, USA, 2003. ACM.

SOUSA v2.0:Automatically Generating Secure and Searchable Data Collection Studies

Brandon L. Kaster*, Emily R. Jacobson+, Walter Moreira*, Brandon Paulson*, Tracy A. Hammond*
Sketch Recognition Lab
Dept. of Comp. Sci. and Engineering
Texas A&M University
Mail Stop 3112
College Station, TX 77843
{bkaster, wmoreira, bpaulson, hammond}@cse.tamu.edu*, em.jacobson@gmail.com+

ABSTRACT

SOUSA is a Sketch-based Online User Study Application developed to aid in the creation of a universal, standardized set of sketch data. This paper describes a secure and searchable interface created for SOUSA (SOUSA v2.0) to make sketch data collection more efficient and practical for researchers and more accessible to a general audience. The expected contribution of our work will be an increase in participation of researchers and practitioners in the field of sketch recognition; thus developing a large, robust repository of sketch data. A motivating factor behind our work is to allow sketch recognition researchers to focus on higher-level tasks, rather than data collection. Features of our interface include a standardized collection mechanism and set of sketch data which will allow new sketch recognition algorithms to be more easily compared with existing models. Our new interface will allow researchers to download and search their own, as well as other publically available, data gathered from collection and verification studies. This new interface will be hosted by the Sketch Recognition Laboratory at Texas A&M University, providing researchers a single, unified solution for sketch data collection and management.

INTRODUCTION

Sketch recognition aims to automatically understand hand-drawn sketched input through an electronic medium [4]. The field of sketch recognition currently lacks a large corpus of hand-drawn data with which to compare new algorithms. Without such a dataset, reported accuracies for new algorithms are not comparable to one another. Although there have been attempts at creating such a corpus, these versions have focused on a few common domains, such as circuit diagrams and family trees, and have not been flexible enough for researchers in more specialized areas [2]. SOUSA v1.0 allowed for users to automatically and simply build web-based data-collection studies for any domain they choose [1]. However, SOUSA v1.0 lacked the ability to give users control over their data; specifically, the system had no way of searching through previously collected data, nor a way of controlling which other researchers can have access to the data they have collected. This paper describes SOUSA v2.0 which allows researchers to easily set up data collection for any domain they choose. Additionally, the system makes it easy to share this data with other researchers, thus creating a repository from which sketch recognition researchers can retrieve sketch data for algorithm testing and comparison purposes.

PREVIOUS WORK

SOUSA v1.0 was created in 2007 by Paulson et al. in the Sketch Recognition Laboratory at Texas A&M University [1]. Designed using a Java interface and Java applet for data collection, this version allows for creation of collection and verification studies. It provided users with a simple management system, but lacked security, and provided data collection study creators no control over who accessed their data. This created a range of problems where users of the system could access any study and edit information that should otherwise be left alone. Lack of such security prevented researchers from creating studies on our server. A list of features is compared between versions of SOUSA in Table 1.

Other corpora of sketch data are limited to specific domains or focus more on labeling primitives in large diagrams. ETCHA Sketches, designed by Oltmans et al. was comprised of diagrams in four domains: circuit diagrams, floor plans, family trees, and basic geometric shapes [2]. Labels were then assigned to the primitives making up each sketch. This work was more focused on labeling portions of sketches and understanding how these primitive shapes are used to create complex diagrams. In contrast, our work is focused on enabling the creation of a large corpus of sketch data in a wide variety of domains and making this data universally available.

A data collection tool was presented at SBIM 2008 by Blagojevic et al. [3]. While this tool also aims to allow researchers to easily collect data in any domain they define, this system focuses more on gathering and labeling large systems and is thus not as flexible for domains in which desired shapes cannot be described in diagrammatic terms. Additionally, our system is accessible through a web interface, making it's studies and data available to researchers worldwide.

IMPLEMENTATION

SOUSA was developed to allow researchers the flexibility to create user studies in any specified domain and to afford researchers a large dataset with which to compare computational styles. Our system allows for both collection and verification studies to be created. SOUSA v2.0 now uses a Python frontend with a MySQL database; JavaScript is used for data collection and is launched from the web interface.

In contrast to SOUSA v1.0, we decided to develop a JavaScript data collector instead of a Java applet to increase the amount of points collected while performing a study. We found that the Java applet recorded on average 80 points/second and Mozilla Firefox v3.0 JavaScript recorded 120 points/second on the same machine.

FRAMEWORK

Management & Security

On our system, users' information and data are contained on a secure database where only users with correct privileges are allowed access to corresponding data and information. Users can change any personal information and have full authority over any studies they have created. The only information stored for a user performing a study is information volunteered for the questions asked by the owner of the corresponding study.

We have provided a search interface to navigate through our current 157 studies and 17,915 data files. This search allows for researchers to quickly find the set of data needed for their research and presents users with a method to quickly find a study to perform. A single user's management of these options can be seen in Figure 5.

Collection Studies

Collection studies can be created in any domain, with the researcher simply specifying labels and/or images depicting the shapes for which she wants to collect samples. Owners of the study control who can perform the study and who can view the corresponding generated data. Questions created are asked prior to a user performing a study and have the option of being deemed mandatory. Shapes have a list of fields to determine the description, number of times each shape is drawn, and the maximum number of strokes for the shape. Each shape also has the option of uploading image(s) to be shown while a user is performing a collection study. An infinite amount of questions and shapes can be added to a single collection study. Figure 1 is the form used to create a collection study.

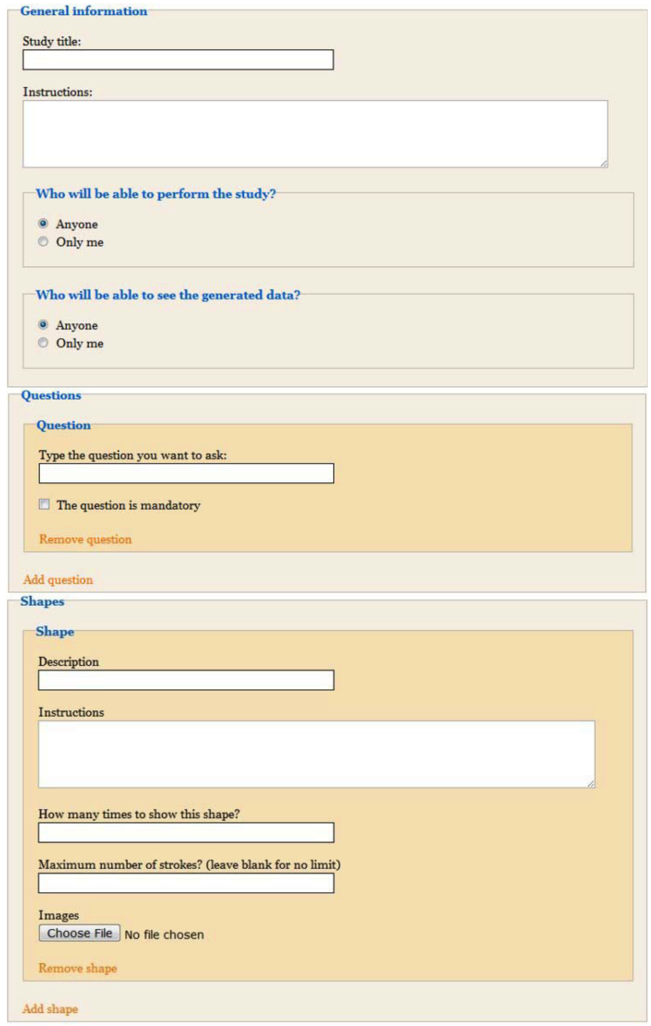


Fig 1. Collection study creation form.

Users performing a collection study are first asked to provide a response to each of the questions asked by the collection study owner. They are then given the option to present the shapes in random or sequential order. For the remainder of the study, users are asked to provide an example of a shape or diagram based on the provided label or image. They are also given a set of options that allows them to undo/redo strokes, suspend/resume studies, clear the panel, and also to continue to the next shape. Figure 2 is a snapshot of a user performing a collection study.

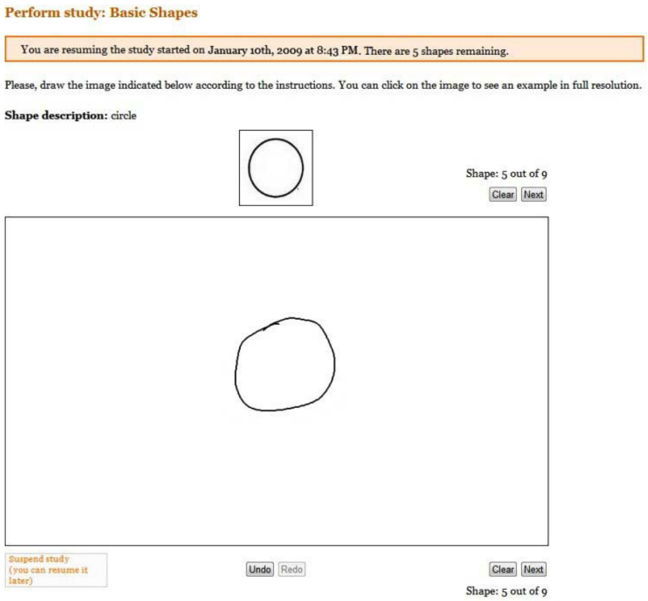


Fig 2. Interface of a user performing a collection study.

Verification Studies

Verification studies are created from an associated collection study and can be used to understand human interpretations of images. Figure 3 shows the form used to create a verification study.

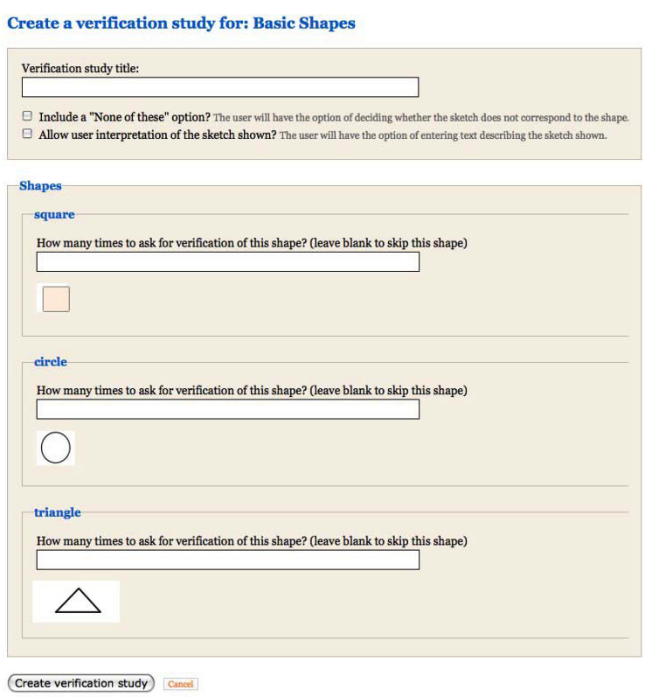


Fig 3. Creating a verification study from an associated collection study.

Performing a verification study is similar to performing a collection study. Users are asked preliminary questions. Users are then shown a drawn shape or diagram and presented with a series of possible labels from which to select. The study participant selects what she thinks the user intended to draw. Figure 4 below shows the user interface seen while taking a verification study.

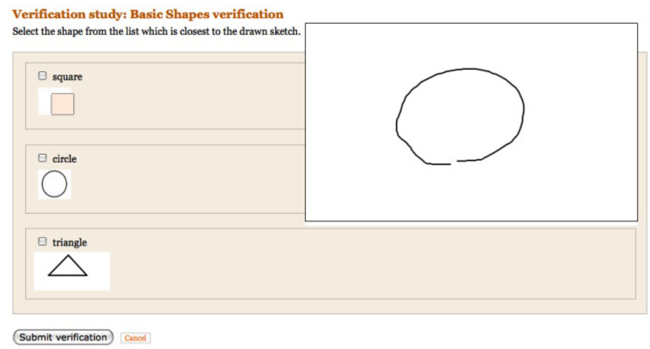


Fig 4. Interface of a user performing a verification study.

FUTURE WORK

The main drawback to our current system is that we lack the ability to archive data associated with an edited or deleted study. If a study is fundamentally changed during an edit, the previous data is associated with the new study and does not reflect the original study from which it was created. Additionally, if a user chooses to delete a study, this data is removed from the system. We hope, in the future, to archive this data, thus maintaining the integrity of the repository while still giving researchers full control over their studies and data.

It would be useful to have a mechanism to save a set of data at a specific moment, so that algorithms being compared could truly use the same set of data, rather than simply data gathered from the same study.

We have also started to work on a deployable miniature SOUSA package for users who must conduct their studies without being connected to the internet. Another idea we would like to implement is to add externally collected data. This would have two purposes: first, to add data previously collected to the repository; second, to allow for the creation of verification studies without an associated collection study.

Further, we would like to develop a mobile platform for mobile data collection. This development provides a collection method to aid research in mobile areas. We have also started development of an API to interact with public data and give users the opportunity to use that data to create other applications of use.

CONTRIBUTIONS

We expect our work to have a broad positive impact on the field of sketch recognition through the creation of a large, robust sketch data repository. Due to the study creation and management of our system, researchers will minimize the time spent collecting sketch data. Since published data is universally accessible, a researcher can utilize previously collected data for her research needs. Most importantly, having a universal set of sketch data in a wide variety of domains will allow new sketch recognition algorithms to be more accurately compared with one another.

ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation (NSF) Grant #0757557 (Pilot: Let your notes come alive) and the National Science Foundation Research Experience for Undergraduates.

REFERENCES

1. B. Paulson, A. Wolin, J. Johnston, and T. Hammond, “SOUSA: Sketch-based Online User Study Applet,” in Proceedings of SBIM, 2008.
2. M. Oltmans, C. Alvarado, and R. Davis , “ETCHA Sketches: Lessons Learned from Collecting Sketch Data”, *Making Pen-Based Interaction Intelligent and Natural*, 2004.
3. R. Blagojevic, B. Plimmer, J. Grundy, and Y. Wang, “A data collection tool for sketch diagrams,” in Proceedings of SBIM, 2008.
4. T. Hammond, B. Eoff, B. Paulson, A. Wolin, K. Dahmen, J. Johnston, P. Rajan, "Free-Sketch Recognition: Putting the CHI in Sketching", in 26th Annual SIGCHI Conference on Human Factors in Computing Systems Works in Progress, 2008.

	SOUSA v1.0	SOUSA v2.0
Server Language	N/A	Python
Client Language	Java	JavaScript
Collection Studies	x	x
Verification Studies	x	x
Dynamically Upload Images	x	x
Secure User Accounts		x
Public or Private access options		x
Browse and Search		x
Download data		x
Zoomed-in images		x
Randomize or sequential display of shapes		x
Start/stop/resume studies		x
Undo/redo in draw panel		x
Support for mandatory or optional questions		x
Classify data, according to criteria, for downloading		x

Table 1. Features compared between released SOUSA versions.

Fig 5. User home page which shows any recently performed unfinished studies and the available options to the owner of a selected study.

Visualizing data to support tracking in food supply chains

Paolo Buono Adalberto L. Simeone Carmelo Ardito
Rosa Lanzilotti
Dipartimento di Informatica - Università di Bari
Via Orabona, 4 - 70125 Bari - Italy
{buono, simeone, ardito, lanzilotti}@di.uniba.it

Abstract

Food tracking has become an important issue in recent times. In light of recent events such as the various outbreaks of food related epidemics, both governmental regulations and increased customer attention, push the demand for systems able to provide knowledge about a supply chain. Tools supporting the traceability and tracking of foods can go a long way towards preventing these unfortunate events or, at the very least, help in minimizing the damage that may occur, thanks to improved and faster intervention possibilities. The study presented in this paper concerns the development of a web application which visualizes supply chain data collected from various companies that are part of the supply chain.

1 Introduction

The problem of tracking and tracing products over their life-cycle is two sided: laws are imposing companies to offer their customers product traceability, while customers want to know what they are exactly buying. Knowing where a product comes from and from which ingredients it is made can contribute to increase the trust between a company and a consumer. Similarly, for organizational and optimization purposes, producers are interested in knowing the steps goods undertake before they become the final retail product.

Actors involved in a supply chain often do not take advantage or cannot access to the overview of a whole supply chain. Verbraeck and van Houten report in their study that "all actors take their own decisions based on limited information, as the other organizations do not share all their data with others" [19]. This is especially true in the supply chain domain that we examined: we interviewed several companies in the food industry and all of them confirmed this truth.

There are benefits in having access to the whole production cycle. Companies might use supply chain data to optimize their production process, while consumers might find interesting to know precisely the steps a product took before ending on her/his dinner table. Due to this dynamic nature of supply chains, it is desirable to organize and structure data through a model [2] and visually present it to the user in graphical form [8]. In fact, information visualization systems have extensively been proven to be beneficial in amplifying human cognition [3]. A web application visualizes an overview of the supply chain allowing users to get details on demand about the visualized data is described in this paper. We advocate that the proposed visualization is helpful both for consumers and for producers.

The paper is organized as follows: in Section 2, an explanation of the supply chain system is given. In Section 3 related work in this field is presented. Section 4 presents the potential stakeholders involved in the system, and the design choices for the graphical layout of the visual representation. A description of the user interface is reported in Section 5. Lastly, in Section 6 we discuss the results from our work and potential future research directions.

2 The supply chain

A supply chain is a network of retailers, distributors, transporters, storage facilities and suppliers that participate in the sale, delivery and production of a particular product. A typical supply chain starts from the gathering of primary resources such as raw ingredients or mineral extractions. These resources are then delivered to the next step in the supply chain, where other companies transform them. This transformation can be a complex process, which entails several steps and involves different companies, each one bringing its contribution until the product is ready to be distributed to retail stores. For example, the beef steaks that we buy at the food store are the result of a long process. First, livestock such as cattle are raised and this requires animal fodder and other kinds of supplies. Successively, the animal is slaughtered and the meat brought to a processing company. In order to manufacture the final product, supplies such as packages, plastic wrap, are needed. This means that supply chains are often intertwined since intermediate goods (e.g. packages, plastic wrap, etc.) needed to refine other products (e.g. beaf steaks) can be themselves the result of other supply chains. Finally, the finished products are then shipped to a distribution company which delivers them to retail stores.

In order to inform about the process that a specific product underwent, a label on the product provides data such as packing date, expiry date, various codes, etc. A code usually found is the production *lot*, that is a reference to the production batch this product was manufactured in; it helps when it is necessary to perform inquires on that specific product. All products that underwent the same transformation and manufacturing processes belong to a same lot. More specifically, a lot refers to a set of products manufactured or packaged in the same conditions. Therefore, products in a lot have the same quality characteristics. A lot is further subdivided in several *pallets*. A pallet is a group of packages which are placed upon a wooden board used to facilitate their transportation.

Currently, individual companies involved in a supply chain do not access information outside their own sphere of influence. The notion of *Supply Chain Management* originated to fulfill the need of integrating key processes characterizing the production cycle [6][11]. By sharing all information available from each company to all parties in the supply chain, both immediate and long term goals can be achieved, such as the improvement of the production process, the optimization of the distribution planning, etc. The gains and the benefits which can be obtained by taking advantage from such an approach can be beneficial to all parties involved in the supply chain and not only to an individual company because the competition focus shifts from a local scale to a global one.

3 Related Work

The main reference about Supply Chain models is the SCOR, *Supply Chain Operations Reference*, developed by the Supply Chain Council (SCC) [13]. SCOR's main goal is to model all the management operations in the Supply Chain and product transactions (both by physical objects and by services) from the supplier's supplier to the customer's customer. It includes methods to gauge delivery and order fulfillment performance, production flexibility, warranty and returns processing costs, inventory and asset turns, and other factors in evaluating the overall effective performance of a supply chain.

Examples of experiences and research on supply chain systems are not very common in literature. SCVisualizer is a prototype system developed to facilitate information sharing within a supply chain [10]. Through the employment of agents making use of webservices it facilitates the retrieval of information from heterogeneous systems in the construction supply chain.

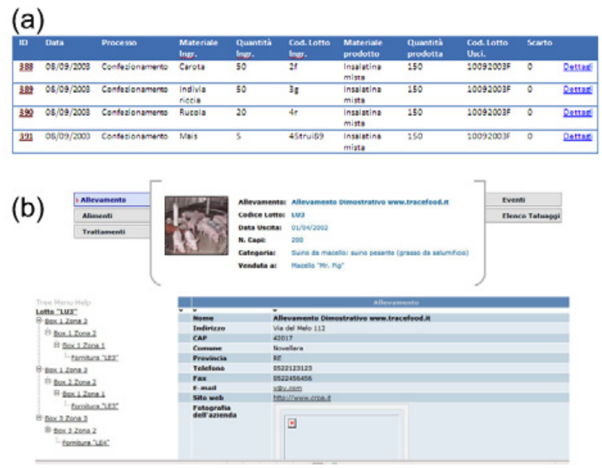


Fig. 1: Examples of the Eti.net web portal (*a*) and the Tracefood.it portal (*b*).

Geotime is a system developed for military personnel that focuses on tracking supply lines from the very beginning to the actual front line [12]. It employs both a pseudo 2D and a fully 3D visualization mode in which the time dimension is modeled in the Z coordinate.

Boyson et al report on the development of a complete portal solution for the F101 turbine engine supply chain [1]. The portal can be used to streamline and automate the workflow operations previously handled via paperwork, such as sharing inventory availabilities between companies, or automatizing spare parts orders, etc.

None of the above systems uses a visualization to represent information: they rely on a geospatial representation of data. ILog LogicNet is a solution aimed at enterprises seeking to analyze and optimize their supply chain [7]. It employs a geographical map together with a Gantt diagram showing the detail of each route.

Among commercially available supply chain portals, various solutions are adopted to visualize production data. Most of these systems present data in textual form, usually as interactive tables which list all production steps with the occasional descriptive image. Each step can therefore be viewed in its entirety, but the context dimension is lost. For example, the Eti.net portal, allows users to visualize data in table form (Fig. 1a) [4]; attributes such as date of the work activity, lot code, etc, are listed. The user can access further details about that particular activity by clicking on an hyperlink placed on each row. This leads to another screen which lists company information related to the entity responsible of performing that particular activity.

Tracefood.it gives users a complete overview of the pro-

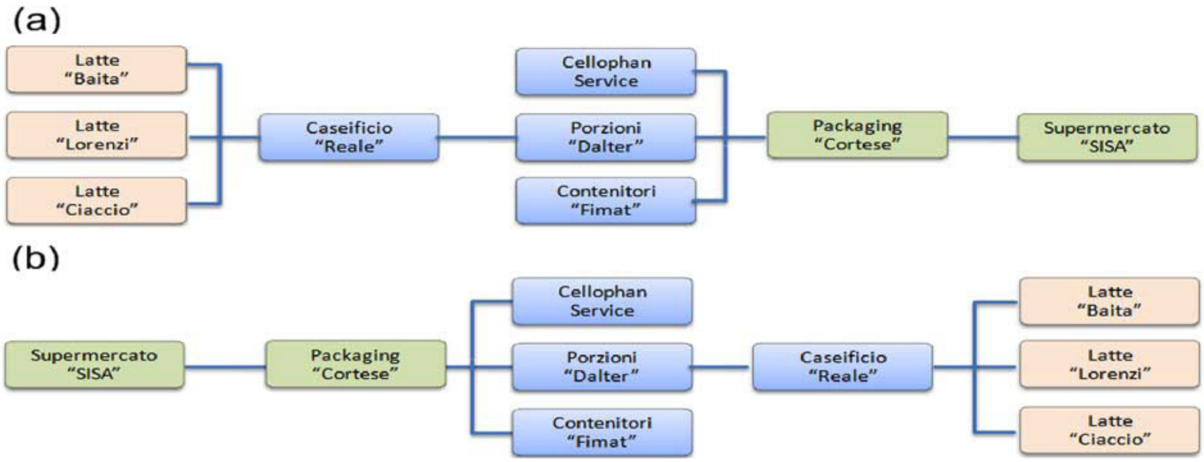


Fig. 2: Two different layouts for (a) tracking or (b) tracing the same supply chain of a piece of cheese.

duction process, down to the last detail (Fig. 1b) [18]. The user can view data about the company involved in the particular production step examined, in text representation. Available data usually is limited to company information and temporal data. Tracefood.it allows users the possibility of querying the database in order to find out where a particular lot was fished. However, it does not permit to view information concerning successive processing.

These visualizations share the possibility of viewing each production step individually. If compared to a visual representation, the textual one gives more prominence to the attributes describing the activities performed during the production process. When the quantity of data increases, by involving several more actors, the supply chain can no longer be expressed in a linear form and instead, it becomes more suited to a tree or graph representations. Furthermore, most of the systems either focus on a single step of the supply chain or they resume more steps in a single one.

Since one approach we considered for the supply chain visualization problem was to use trees or hierarchical graphs, we surveyed some known approach in literature [16].

Treeplus is a tool for visualizing graphs as trees [9]. It places nodes according to their depth level and, after focusing in a node, it highlights connected nodes in order to give users an understanding of the graph structure. There are also systems that employ a temporal classification of nodes.

Netscan is a tool designed to explore newsgroups threads data [15]. One of its visualizations uses a tree representation where its nodes (which represent messages) are grouped in vertical bands depending on the calendar days the messages in the thread are posted. Edges between messages mean that the more recent one is a reply to the older one.

Treevolution is another general purpose tool that uses an innovative approach to visualize trees. All the nodes are grouped according to their temporal date and placed on a series of concentric rings around the root, imitating a tree ring metaphor [17].

4 Designing a visual representation

Although most of the related work on supply chains focuses on the geographical representation of involved entities (as far as graphical visualization is concerned), our focus is instead on the different stages that a product undergoes in a supply chain, from supplier to consumer or viceversa.

Through on site interviews with company managers and other people interested in this topic, we identified the main potential stakeholders: producers, supply chain operators, company managers, public agencies, sellers, authorities, consumers.

Producers provide the raw materials to the production/transformation companies.

Supply chain operators include people directly responsible with the internal management of the supply chain that work in a company typically involved in the manufacturing phase. These people are interested in ways to optimize their production processes by taking advantage of the readily available information about their suppliers and distribution partners, that such a visualization system can supply.

Company managers are responsible of the quality of the finished products of the supply chain nodes.

Public agencies guarantee the quality of foodstuffs. They also perform quality tests, e.g. those required by the H.A.C.C.P. (*Hazard Analysis and Critical Control Points*) European rules which define key actions that can be taken to reduce or eliminate the risk of hazards.

Authorities are public agencies that have the responsibility of controlling products available in the market. They are clearly interested in any potential event that might en-

Fig. 3: The user interface of the web application that shows the example of tracing a piece of cheese.

danger people. Assuming that some event causes inadequacies in the production process at a certain facility (livestock living conditions, health conditions, etc), authorities should promptly recall whole batches of products originated from that facility.

Sellers are supply chain operators with the responsibility of selling products to the final user, the consumer.

Consumers are the final users that buy the product at the end of the supply chain. They may find interesting knowing the production history behind a product they bought.

The main entities that can be displayed are grouped into four main categories: suppliers of raw elements, manufacturers, distributors and retail points. At a first glance, the simplest representation seems to be the tree one, in which the father node represents a company that supplies a good that will be used in the next production step by another company. Eventually we ended up using a hierarchical graph representation, for the following reasons.

In a tree representation, there can only be one root and each node can only have one father. But, sometimes the tree representation is not suited to show the facts that happen in the supply chain domain. In fact, in the example shown in Fig. 2, the ”Caseificio Reale” dairy, to produce its cheese, needs milk from three different producers (”Baita”, ”Lorenzi” and ”Ciaccio”). In this case, the ”Caseificio Reale” node has three fathers and we cannot represent this situation using a tree.

In the supply chain domain, the temporal order in which production steps are performed (e.g. cheese production requires milk to be produced first) is another important data characteristic that should be preserved. Certain nodes could have no parent nodes, due to them entering the production process at later stages. For example, in Fig. 2 the ”Contenitori Fimat” company that supplies containers does not have any father node, because its containers are not needed until the cheese is ready to be packaged in its final retail form, after several months of maturation. One possible way to overcome this problem is to arbitrarily choose a root node and reorganize the other nodes accordingly. But, if in Fig. 2 *(a)* we assume that the ”Caseificio Reale”' node is to become the root of the tree, the three milk producers, being connected to it, would become its children. In this way, readers could wrongly infer from the representation that the milk production happened after the cheese production.

In order to accommodate complex cases that can happen in the supply chain domain, a graph structure is needed. We decided in favor of a hierarchical graph representation in which edges connecting nodes represent the transportation mean through which a good was moved from one location to another. In order to increase its readability, we applied some constraints on the graph representation. Namely, the nodes that represent steps performed at the same temporal stage of the supply chain are aligned in the same column.

By assuming a left-to-right reading order, there are two possible orders in which to place the elements, depending on whether to consider the production process from producer to retailer or viceversa. Each of the two different designs (Fig. 2) is better suited to highlight different aspects of a supply chain. In fact, by placing the starting steps of the production process (the gathering of the ingredients from suppliers, for example) on the left part of the diagram the accent is placed on tracking the subsequent steps as the raw resources get progressively transformed into the finished good (diagram *a* in Fig. 2). On the contrary, by placing the final steps on the left side of the screen, the accent is instead placed on product traceability (diagram *b* in Fig. 2). It is in fact easier to retrace and follow the history of a product to its origins.

5 Tracking a good in a supply chain

From the analysis of related work, users and several interviews to various actors of the supply chain, we gath-

Fig. 4: A close-up of a node.

ered important information about relevant tasks for people that want to use this application. Among various tasks, we present here some examples to justify our proposal. Given a product, a user may be interested in: 1) visualizing the whole supply chain of that product; 2) getting information about a company producing a specific ingredient of a product; 3) knowing the time to transport some goods from different locations in the supply chain.

Let us suppose that a person (female) is interested in knowing when and how a parmesan cheese piece, bought in a "Supermercato Sisa" store, was actually shipped out from its manufacturer. She has to connect to the web portal (our tool has been developed as a Silverlight web application [14]). She is presented with a web form where the lot code (that can be found on the cheese package) must be inputted. Then the visualization shown in Fig. 3 appears.

We used a hierarchical graph representation in which each node represents a step of the production process that falls in one of these categories: facilities that supply a raw good, transform it, distribute it or sell it to the public. Each node of the graph is placed from left to right according to their temporal stage. Each node (see Fig. 4) is characterized by the the name *(a)* of the company, the logo *(b)*, a label indicating the production stage in the supply chain *(c)*, two buttons used respectively for expanding and collapsing the graph (the number in button (*d*) refers to all children of that node, while the number in button (*e*) refer to all its descendants) and a button used for hiding the node and its children in the graph *(f)*, labeled with the usual "X" symbol that closes a window. The visualized nodes are connected to each other by means of edges which symbolize a physical transportation of the good from one location to another. The iconic label on the edge indicates the transportation mean (e.g. a wheel for a truck). By clicking on it, further details can be examined, including the time taken for the transportation. In the example in Fig. 3, the parmesan cheese piece that the user is tracing has been sold in the "Supermercato SISA" (leftmost node in the graph). It was shipped out by a truck of the "Cortese Packaging" company (previous step), which packaged the final product. To accomplish this work, "Cortese Packaging" receives the parmesan cheese pieces by "Porzioni Dalter", the plastic wrap by "Cellophan Service" and the container by "Contenitori Fimat" (step 3 in Fig. 3). The supply chain visualization continues until the companies that produced the animal fodder needed by the cattle that produced the milk from which the parmesan cheese has been made.

Fig. 5: Detailed information about a node.

By clicking on a node, a panel is displayed which shows detailed information about that node. In the example of Fig. 5, our user asked for details about the "Caseificio Reale" that produced the whole parmesan cheese wheel. In the upper part of the panel, company address and contact data are shown; in the bottom part of the panel are details about the lot code and production date along with other information.

In order to help users notice any change in the graph representation, animations are used whenever the graph has to reorganize in a different layout, from the old positions to the new ones. A slider control is available to change the zoom ratio (in the top left part of Fig. 3), allowing users to zoom out so to fit the whole graph into the screen or zoom into a close up of a specific part of it.

A search feature is available to allow the user to quickly find a certain node in the graph. She can type a text string in the searchbox placed in the top right part of the interface shown in Fig. 3. The string will be searched either in the nodes' label or in the other fields. Nodes which satisfy the search constraints are highlighted, together with the path needed to reach it from the starting node.

In order to shift the focus from traceability to tracking of a product or viceversa, a special *inverse* command is available (in the first button in the top left part of Fig. 3) to mirror the graph by bringing the rightmost nodes to the left side of the screen and viceversa.

6 Conclusion

In this paper we have presented a diagrammatic representation of supply chains focusing on the food industry. We have provided several features that allow users to interact with this representation and get information not only about products and their ingredients, but also about companies that produce raw materials. Government regulations in most countries require companies to document the various steps their products undergo. Documents and data needed to accomplish the tasks we outlined for the different user categories are already there, although the knowledge on which this kind of analysis is possible is mostly available only on paper. However, knowledge outside the companies' own zone of competence is severely lacking because companies are usually only interested in what happens inside their own facilities. The growing complexity of supply chains require an increased awareness of the external parties involved, in order to comply with stricter government regulations, increasing users' requests, and to maximize production performance.

In light of this, visualizing information that previously existed only as paper documents gives users the unique advantage of allowing them to focus on parts of the data to be studied more carefully [5]. Compared to websites that offer only textual representations, thus giving users a very limited and simplified view of a vastly more complex process than what they can observe, the system prototype we created enables users to access, through a simple interface, a complete view of every aspect involved in the food supply chain: from the production stage, until the moment where the product is delivered to retail stores. Details and interesting insights which could be hidden in the data can now be more easily identified thanks to the visual interaction features at the users' disposal.

This is an ongoing work and we are planning to expand the users' querying possibilities in order to provide them with better filtering mechanisms and reporting views. We are also planning to perform user studies in real settings, to evaluate the impact of the use of graphical visualization to assist users in their work. As a future research direction, we are going to focus on new ways in which both textual and graphical information can be integrated to facilitate users in gathering the required information.

Acknowledgment

Partial support for this research was provided by EU and Regione Puglia under grant DIPIS. We thank Prof. Maria Francesca Costabile for her valuable guidance in our research and Dr. Catherine Plaisant for the stimulating discussions about visualization techniques.

References

[1] S. Boyson, T. Corsi, and A. Verbraeck. The e-supply chain portal: a core business model. *Transportation Research Part E: Logistics and Transportation Review*, 39(2):175 – 192, 2003.

[2] S. Boyson, L. H. Harrington, and T. M. Corsi. *In Real Time: Managing the New Supply Chain*. Praeger Publishers, 2004.

[3] S. K. Card, J. D. Mackinlay, and B. Shneiderman, editors. *Readings in information visualization: using vision to think*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1999.

[4] Eti.net, http://www.etinet.biz/. Last accessed 30/04/2009.

[5] J.-D. Fekete, J. J. Wijk, J. T. Stasko, and C. North. *The Value of Information Visualization*. Springer-Verlag, Berlin, Heidelberg, 2008.

[6] C. M. Harland. Supply chain management: Relationships, chains and networks. *British Journal of Management*, 7(s1):S63 – S80, 1996.

[7] ILog Logic Net, http://www.ilog.com/products/logicnet-plus-xe/. Last accessed 30/04/2009.

[8] C. Kilger and B. Reuter. Collaborative planning. In H. Stadtler and C. Kilger, editors, *Supply Chain Management and Advanced Planning*, pages 259–278. Springer, Berlin, 2 edition edition, 2002.

[9] B. Lee, C. S. Parr, C. Plaisant, B. B. Bederson, V. D. Veksler, W. D. Gray, and C. Kotfila. Treeplus: Interactive exploration of networks with enhanced tree layouts. *IEEE Transactions on Visualization and Computer Graphics*, 12(6):1414–1426, 2006.

[10] J. U. Min, H. Bjornsson, J. U. Min, and H. C. Bjornsson. Construction supply chain visualization through web services integration. Technical report, Center for Integrated Facility Engineering, University of Stanford, 2004.

[11] R. K. Oliver and M. D. Webber. Supply-chain management: logistics catches up with strategy. *Logistics: The Strategic Issues*, pages 63–75, 1992.

[12] P. Proulx, L. Chien, R. Harper, D. Schroh, D. Kapler, D. Jonker, and W. Wright. nspace and geotime: A vast 2006 case study. *Computer Graphics and Applications, IEEE*, 27(5):46–56, Sept.-Oct. 2007.

[13] SCOR, http://www.supply-chain.org/. Last accessed 30/04/2009.

[14] Silverlight, http://silverlight.net/. Last accessed 30/04/2009.

[15] M. A. Smith and A. T. Fiore. Visualization components for persistent conversations. In *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 136–143, New York, NY, USA, 2001. ACM.

[16] K. Sugiyama, S. Tagawa, and M. Toda. Methods for visual understanding of hierarchical system structures. *Systems, Man and Cybernetics, IEEE Transactions on*, 11(2):109–125, Feb. 1981.

[17] R. Therón. Hierarchical-temporal data visualization using a tree-ring metaphor. In *Smart Graphics*, pages 70–81, 2006.

[18] Tracefood.it, http://www.tracefood.it/. Last accessed 30/04/2009.

[19] A. Verbraeck and S.-P. A. van Houten. From simulation to gaming: an object-oriented supply chain training library. In *WSC '05: Proceedings of the 37th conference on Winter simulation*, pages 2346–2354. Winter Simulation Conference, 2005.

A methodological framework for automatic clutter reduction in Visual Analytics

Enrico Bertini [1] and Giuseppe Santucci [2]
[1] DIUF/DIVA - University of Fribourg Bd de Prolles, 90
CH 1700 Fribourg, Switzerland - enrico.bertini@unifr.ch
[2] Dipartimento di Informatica e Sistemistica - Sapienza Università di Roma
Via Ariosto 25 - 00185 Rome Italy - santucci@dis.uniroma1.it

Abstract

Visual Analitycs manages billions of data records and, in such a case, pure Information Visualization reveals its limits in terms of expressivity and scalability. In this paper we deal with the critical issue of visual clutter reduction, presenting a general framework for analyzing and reducing clutter in infovis techniques. We detail our proposal using two running examples, i.e., 2D scatterplots and parallel coordinates.

1 Introduction

Visual Analytics (VA) presents several research challenges and a recent dedicated workshop [1] pointed out the most relevant ones. Among them, we deal with the issue of integrating automated data analysis methods to reduce image clutter.

In the last years a large number of clutter reduction techniques have been proposed (a survey is in [8]); in this paper, we present a methodological framework able to deal, in a general way, with the issue of cluttered images. The framework foresees four steps and uses quality metrics and user studies. In the paper we use two running examples, i.e., 2D scatterplots and parallel coordinates; we believe, however, that the proposed approach is general enough to be applied to many others visualization techniques.

The paper is structured as follows: Section 2 describes the framework ideas and steps, Section 3 analyzes related works, Section 4, 5, 6, and 7 describe the application of the methodological steps to the two running examples, 2D scatterplot and parallel coordinates. Section 8 presents conclusions and future work.

2 The methodological framework

Our approach is based on the idea of modeling clutter in a formal way, to have some figures that can be (a) validated against human perception and (b) used to drive/validate a clutter reduction algorithm. The focus of the paper is not on defining a new strategy but on proposing several steps that allow for using a clutter reduction approach in a more sound way. In particular we foresee four methodological steps:

1. Model clutter in the target visualization. The goal of this step is to define basic measures and formal models able to characterize *numerically* the clutter. Because we are dealing with digital images, some measures are common to every visualization (but the methods to derive them could be very different); on the other hand a particular visualization can be characterized by measures that are specific for it.

2. Validate measures against human perception. Figures are mandatory (you cannot control what you cannot measure) but human perception does not follows mathematics in a straightforward way, so it is needed to understand how the figures we are dealing with are perceived by human eyes.

3. Define Quality Metrics. Composing basic measures allows for devising more complete figures about the overall image quality.

4. Use Quality Metrics. User validated figure are an objective way for validating/driving an automatic clutter reduction technique.

3 Related work

The main objective of our investigation is to model the quality of a visualization trhough measurable parameters and to use them to drive optimization algorithms. We proposed this idea [3], where some initial clutter metrics and uniform sampling on scatterplots were used. Since then, the same basic idea was expanded and refined with additional metrics, visualization techniques, and recovery algorithms [4].

A similar goal has been pursued by other researcher whose main goal is to find the best mapping between the dataset and specific visual configurations. Scagnostics is a graph-theoretic technique first introduced by Tukey [12] and recently expanded by Wilkinson et al. [13] in which the idea is to compute a series of measures on a scatterplot matrix, which maps all the data dimensions in a series of 2D scatterplots, to select pairs of dimensions showing the most interesting patterns. A similar approach is used in [11] where the same basic framework is used and extended to include pixel-based visualizations. Another similar proposal is in [2] where a perception-based search algorithm is employed to find the best mapping between data and visual features. A search in this database is performed through data mining techniques in order to find optimal solutions to a given visualization problem.

In our work the focus is shifted from the problem of finding an optimal mapping to the problem of optimizing a given visualization when the mapping is already given. The same approach is adopted by Peng et al. in [9] where a standard set of nD visualizations is provided (e.g., parallel coordinated, scatterplot matrices, etc.). In this work each visualization is modeled in terms of the kind of clutter it generates and optimized through various techniques like: axes reordering, sampling and aggregation. In [6] a step further is performed by implementing a method to provide to the user a feedback on the current quality/data loss of the visualization when interactively changing the level of data abstraction through aggregation or sampling.

Ellis and Dix propose another similar study in [8] where only parallel coordinates are taken into account. The authors propose a series of statistical models to estimate the clutter level obtained when increasing the number of displayed data items. The technique is then used to drive a fast algorithm that selects the right amount of sampling needed to see the details within a "‘sampling lens"’, that is, a movable lens within which sampling is applied.

3.1 Positioning our approach

Our approach differs from the discussed proposals because instead of focusing on a specific visualization and clutter reduction method it proposes a methodological framework based on clutter models, metrics, and user studies.

Concerning the running examples we use in the paper we point out some unique features of our decluttering techniques, described in the following.

- Measuring lost features. Our quality metrics for 2D scatterplots allow for discovering, in a quantitative way, whether relevant data characteristics are preserved in the actual visualization;

- Providing a detailed overview without altering the image size. Zooming, displacement, sampling, and density maps are commonly used to reduce clutter in 2D scatterplots. Displacing elements introduces some errors in the representation, while zooming causes to lose the overall image perception, unless the zoom is applied to the whole image (in this case, the screen dimension represents the limit of this approach). Sampling and density maps may effectively improve the image readability without enlarging it; however, if the data present quite different density values these techniques are not very effective: faint zones may disappear while making the most dense areas readable and little density differences are hardly perceivable by the user. Our nonuniform sampling technique, described in Section 6.1, addresses exactly this problem, providing a data overview preserving faint zones and showing details in denser areas, without altering the image size.

- Parallel coordinates clutter characterization. The paper present a novel statistical approach for estimating parallel coordinates clutter. The model allows for defining some general quality metrics that have been refined through user studies and used to drive automatic sampling techniques.

A work similar to our proposal is the one by Schneidewind, Sips, and Keim in [11]: it foresees a three steps approach and uses data analysis and measures at pixel level to automatically select the most relevant attributes and images to present to the end user. On the other hand, that strategy is quite far from our method: we work on a *unique* image, measuring collisions and plotted pixels to reduce the clutter and we measure both the data and the visualization spaces, comparing the figures to derive quality metrics.

4 Step 1: Modeling Clutter

In this section we present a statistical framework characterizing the basic figures that allows for modeling a cluttered visualization.

4.1 Modeling Clutter in 2D scatterplots

A complete model for characterizing the clutter of 2D scatterplot has been presented in [4]; here we recall the results useful for detailing the methodological framework. We consider a 2D space in which we plot items by associating a pixel to each data element and the pixel position is computed mapping two data attributes on the spatial coordinates. We derived a a series of functions that estimate the amount of colliding points and, as a consequence, the amount of free available space. We used such functions to estimate data density, splitting the whole image in squares of area A, say *sample areas* (SA) and we calculate two different densities: *data density* and *represented density*.

Data density is defined as $D_{i,j} = \frac{n_{i,j}}{A}$ where $n_{i,j}$ is the number of data points that fall into sample area $A_{i,j}$. If we plot n data elements, each $SA_{i,j}$ assumes a value within the set $0, \frac{1}{A}, \frac{2}{A}, \ldots, \frac{n}{A}$. For each distinct value we can count the number of sample areas characterized by that value, obtaining the data density distribution.

Represented density is defined as $RD_{i,j} = \frac{p_{i,j}}{A}$ where $p_{i,j}$ is the number of distinct active pixels in $SA_{i,j}$. The number of different values that a represented density can assume depends on the size of sample areas. If we adopt sample areas of 8×8 pixels the number of different not null represented densities is 64.

Because of collisions the number of active pixels on a sample area $SA_{i,j}$ will likely be less than the plotted points so $RD_{i,j} \leq D_{i,j}$.

4.2 Modeling Clutter in Parallel Coordinates

Measuring plotted pixels and collisions in the parallel coordinate context is not straightforward and few attempts are available in the literature. The main problem is that each data item is plotted on a *varying* number of pixels and that the collision between two lines involves again a *varying* number of pixels. An approach very close to the one presented in this paper is in [7] where clutter is computed within a prefixed area (i.e., a sampling lens) using a probabilistic method. We improve that result estimating the average number of pixels plotted for each data item and the average number of pixels involved in a single collision. Using these figures we are able to estimate the number of plotted pixels and collisions.

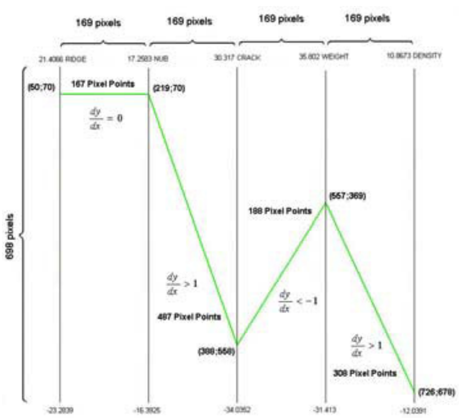


Figure 1. Varying number of pixels.

It is quite evident that the number of used pixels M by a line depends on the line slope and on the way lines are drawn on discrete pixels (see Figure 1). In the following we concentrate on a pair l of adjacent attributes displayed on $x \times y$ pixels. The shortest line we can draw (horizontal line) measures x pixels; the longest one (diagonal line) measures y pixels. The overall number of distinct lines is y^2 and considering the aleatory variable X_l representing the length of a line drawn across the pair i we have that $\mu_l = E_l(X_l)$ gives us the information on the average length. We computed the frequency distribution of M, and the corresponding probability function $p_l(M)$.

If N_T is the number of data items plotted on the screen the expected values of plotted pixels is $\mu_l \times N_T$. Note that this value has the same meaning as the n value we used in the 2D scatterplot case: it represents the number of pixels the data are going to plot. In the case of 2D scatterplot it corresponds to the data items; here we needed an intermediate calculation.

In order to compute the overall number of collisions, we first focus on the expected number of collision between two lines, $E_l(C)$. We compute this value on a pair l of adjacent attributes, comparing all the possible line couples and counting the number of colliding pixels. Using this figures we compute the frequency distribution and the corresponding probability function. The problem is that, in such a case, the cost of the computation is very high $O(y^4)$ and it is not feasible in real cases (response time>10 hours). We tried to estimate this value using only a subset of all $O(y^4)$ pairs but, in order to have a precision better than 1%, we come up with a system still unusable (response time>20 seconds). Eventually, we decided to construct an interpolation function using 520 x,y samples in the interval [10,1000], [50, 1000] with a precision better than 0.25%. It is worth noting that while x and y are similar $E_l(C)$ presents low values in the interval $[0.7 \ldots 4]$. As one of the axis is much greater than the other one (green/red zones) the expected number of collisions between two lines increases quickly.

Once we know the $E_l(C)$ value, we can compute the overall number of collisions through the following formula:

$$TotC_l = \sum_{i=2}^{N_T} min[\mu_l; \mu_l - \mu_l(1 - \frac{E_l(C)}{\mu_l})^{i-1}]$$

We consider that when a new line is drawn the maximum number of colliding points is μ_l and we estimated the actual number of collisions computing the number of μ_l points that likely are not involved in collisions using the expression $(1 - \frac{E_l(C)}{\mu_l})^{i-1}$ that corresponds to the probability that a single pixel of the i_{th} line (among μ_l) is *not* involved in a collision with the already drawn i-1 lines. A set of simulations confirmed us the validity of the model we are using.

5 Step 2 : Perceptual validation of metrics

This section describes the studies we performed to understand the relationships between basic metrics describing the clutter and user perception of visual clues.

5.1 2D scatterplots

Among the measures we defined in Section 4.1 we investigate the way in which *represented densities* are perceived by human beings, computing what is the minimum difference in active pixels between two sample areas that allows for perceiving a density difference.

To answer this question we performed a study [4] based on a comparison strategy [10], asking the users to recognize few more dense areas on a uniform background (basis), repeating the test for different bases and different density differences. Linearly interpolating these values we derived a function $minimum\delta(RD_{i,j})$ returning the *minimum* increment a sample area must show to be perceived as denser than $SA_{i,j}$

5.2 Parallel coordinates

Concerning parallel coordinates we investigated how clutter hides visual clues. In particular, we set up different data sets presenting strong visual relationships between two attributes, i.e., clusters of lines and we hide these relationships in very cluttered images. The idea was to ameliorate the image in several steps (e.g., sampling the data set with increasing sample ratios or enlarge in several steps the drawing area) till the user was able to recognize the patterns.

We involved 45 people (24 males and 21 females, ranging between 20 to 65) using six different data sets and we averaged their answers in terms of number of displayed elements N_T, number of plotted points n, and number of collisions $TotC_l$ corresponding to the situation in which they were able to recognize the patterns. Figure 2 presents some sampling steps presented to the users in order to recognize a pattern. The yellow color denotes that the user recognized the patterns. The collected figures allow for defining useful threshold values for the quality metrics discussed in the next section.

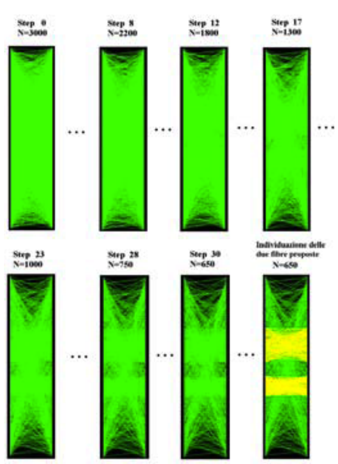


Figure 2. Recognizing patterns

6 Step 3: Defining Quality Metrics

In this section we provide several quality metrics, intended for measuring the image degradation and we discuss the way in which they can be used. A comprehensive discussion of Quality Metrics typologies and their usage is out of the scope of this paper; more details are available in [5].

6.1 Quality Metrics in 2D scatterplots

A broader description of 2D scatterplot quality metrics has been presented in [4]; here we recall the main results. Starting from the basic measures we described in Section 4.1 we define some quality metrics in order to evaluate how much we are overplotting the screen, the collision percentage, and how many density differences are still visible in the displayed image.

The first quality metric we provide is the following:

$$SP\_PPr\text{(scatterplot\_Points-Pixels ratio)}=\frac{n}{x \times y}$$

that measures how much we are overplotting the screen. After that we measure the percentage of colliding points:

$$SP\_CPr\text{(scatterplot\_Collisions-Points ratio)}=\frac{k}{n}$$

that measures how many data items are colliding.

These two quality metrics provide a quick and effective information about the image clutter and are quite general, i.e., they do apply to any visualization.

If we want to explore specific 2D scatterplot quality metrics we consider relative densities and we measure the lost density differences through the metric PLDDr (Perceptually Lost Data Densities ratio). This quality metric is computed comparing couples of sample areas and checking whether their relative data density (D) is *visually* preserved or not when considering their represented density (RD). To compute such a metric we designed an algorithm that considers all the couples of SAs, comparing their data density (D) and represented density (RD) and counting the non-matching pairs.

6.2 Quality Metrics in Parallel Coordinates

Starting from the basic measures we described in Section 4.2 we define, analogously to the 2D scatterplot case, two quality metrics to evaluate how much we are overplotting the screen and what is the collision percentage.

The first quality metric we provide is the following:

$$PC\_PPr \text{ (Parallel Coordinates\_Points-Pixel ratio)}= \frac{\mu_l N_T}{x \times y}$$

In this case we need to use the statistical results we devised Section 4.2 : drawing N_T lines on a pair l displayed on $x \times y$ pixels corresponds to plotting $n = \mu \times N_T$ data items on a $x \times y$ area. The results of the user study allowed us to compute a threshold value TV_{PC_PPr} of 3.24.

Roughly speaking, this value indicates that when we are drawing a number of lines that require more than three times (i.e., 3.24) the available pixels the relationships between attribute pairs are not perceptible anymore.

The second quality metric measures the percentage of colliding plotted pixels:

PC_CPPr =(PC_Collision-Plotted Point ratio)= $\frac{TotC_l}{\mu_l N_T}$

For this quality metric the user study produced a threshold value TV_{PC_CPPr} of 0.73. Roughly speaking, this value indicates that when the number of collisions is above the 73% relationships between attribute pairs are not perceptible anymore.

7 Step 4: Using Quality Metrics

In this section we show the role that quality metrics can play in clutter reduction; in particular, we foresee three main usage patterns:

1. given a data set and an ameliorating algorithm, it is possible use a quality metric to validate its effectiveness;
2. given a data set and a prefixed display area size, it is possible to check a metric(s) (or a linear combination of two or more quality metrics) against a threshold value(s) and, in case of violation, to sample the data until the threshold(s) is satisfied (or looking for the sampling ratio that optimizes the metric if it is nonmonotonic);
3. given a data set and a set of threshold values, it is possible to compute the *minimum* area size preserving them, allowing the system to optimize the screen usage.

7.1 Validating a 2D scatterplot clutter reduction algorithm

Several proposals attempt to reduce clutter applying the same amount of sampling to the whole image. While this is quite straightforward it presents several drawbacks. For instance, to sample areas presenting very low data density is useless and potentially dangerous, because empty areas may appear where data was previously plotted. In addition, in some cases the user is interested in discovering as many density differences as possible, neglecting their intensity. Therefore, we devised a non uniform sampling approach able to (a) preserve low density areas and (b) to show to the user a greater number of density differences. Details about the algorithm are in in [4]; here we recall the main algorithm idea and we show how the PLDDr() quality metric allows for *validating* the result in an objective way.

We apply our techniques against the dataset shown on, Figure 3, containing about 160,000 mail parcels plotted on the X-Y plane according to their weight (X axis) and volume (Y axis). The image is displayed on a 304×304 screen and using 8×8 pixels SAs it is characterized by the following quality metrics: SP_CPPr=0.79 and SP_CPr=0.80; even if the occupation of the screen is very little (we are plotting less points than pixels), the area close to the origin is very crowded and presents a great number of collisions (80% of the data set is colliding).

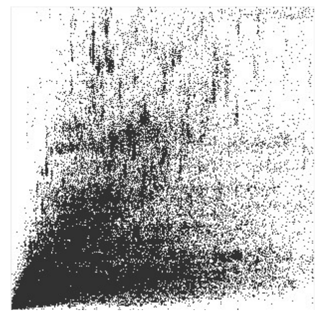


Figure 3. Original image (304 x 304 pixels)

The main idea of our nonuniform algorithm is to apply a *different* sample ratio to each sample area, to obtain a represented density distribution as uniform as possible, increasing the number of different represented densities. In fact, the problem of preserving relative densities can be challenged altering the mapping between the set of the actual data densities and the set of available represented densities; while the uniform sampling approach alters the mapping in linear way, nonuniform sampling forces different data densities to be represented on the *same* represented density, through different sampling ratios.

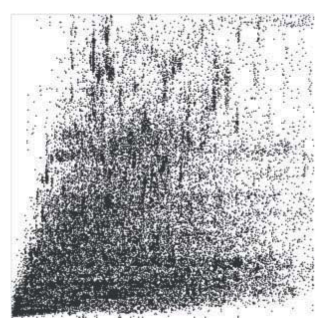


Figure 4. nonuniform sampling

Figure 4 shows the result obtained applying the nonuniform sampling algorithm. High density areas present more density differences than the original image. The PLDDr metric validates the visual impression: it is equal to 0.43 against the original value 0.63 : the algorithm rescued more than 30% lost density differences. Moreover, the algorithm does not alter low density areas that are represented as they appear on the original image.

7.2 Driving sampling in parallel coordinates

In this case we apply a uniform sampling algorithm, sampling the data till a threshold value is satisfied.

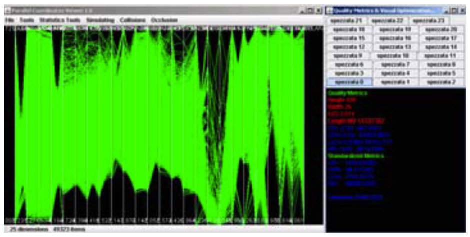


Figure 5. Crowded Botanical data

We consider the crowded dataset in Figure 5, 49323 25-dimensional botanical data items. Each adjacent pair of attributes is displayed on a 430×26 pixels area and the image exhibits very poor quality metrics: PC_PPR=687.6, meaning that we are plotting a number of pixels 687 times greater than the number of available pixels (!) and PC_CPPR=0.99975 meaning that 99.9 % of the plotted pixels are colliding (!). We select the PC_PPR metric to drive the sampling algorithm: data is sampled till PC_PPR is below the threshold value (3.24). The result is presented in figure Figure 6 containing the 0.45 % of the original data set: most of the visual relationships between adjacent attributes are now visible. It is worth noting that the metric PC_CPPR decreased till 0.77, a value very close to the threshold (0.73), confirming the threshold values we identified through the user study.

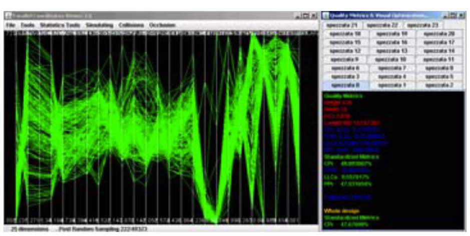


Figure 6. Threshold driven sampling

8 Conclusion And Future Work

In this paper we presented a methodological framework for automatically reducing clutter in VA. The framework uses basic measures, user studies, quality metrics, and clutter reduction algorithms.

The methodological steps have been demonstrated by means of two widely used visualization techniques, 2D scatterplots and parallel coordinates.

We are actually extending our approach to other visual techniques and exploring a wider use of quality metrics in the VA context.

References

[1] Workshop on Visual Analytics, Darmstadt, Germany, June 2005 http://infovis.uni-konstanz.de/index.php?region=events&event=VisAnalyticsWs05.

[2] Alethea S. Bair and Colin Ware. An approach to the perceptual optimization of complex visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 12(4):509–521, 2006. Member-Donald H. House.

[3] Enrico Bertini and Giuseppe Santucci. Quality metrics for 2d scatterplot graphics: automatically reducing visual clutter. In *Proc. of 4th International Symposium on SmartGraphics*, May 2004.

[4] Enrico Bertini and Giuseppe Santucci. Give chance a chance: modeling density to enhance scatter plot quality through random data sampling. *Information visualization*, 5(2):95–110, May 2006.

[5] Enrico Bertini and Giuseppe Santucci. Visual quality metrics. In *Proceedings of BELIV'06 BEyond time and errors: novel evaLuation methods for Information Visualization - International workshop of the AVI 2006 International Working Conference*. ACM Digital Library, 2006.

[6] Qingguang Cui and Jing Yang. Measuring data abstraction quality in multiresolution visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):709–716, 2006. Member-Matthew Ward and Member-Elke Rundensteiner.

[7] G.Ellis and A.Dix. Enabling automatic clutter reduction in parallel coordinates plots. In *Proceedings of IEEE Visualization*, November 2006.

[8] G.Ellis and A.Dix. A taxonomy of clutter reduction for information visualization. In *Proceedings of IEEE Visualization*, October 2007.

[9] Wei Peng, Matthew O. Ward, and Elke A. Rundensteiner. Clutter reduction in multi-dimensional data visualization using dimension reordering. In *Proc. of IEEE Symposium on Information Visualization*, pages 89–96. IEEE Computer Society, 2004.

[10] F. Vardabasso S. Alam, R. Luccio. Regularity, exposure time and perception of numerosity. *Perceptual and motor skills*, 63(2), 1986.

[11] J. Schneidewind, M. Sips, and D.A. Keim. Pixnostics: Towards measuring the value of visualization. In *Visual Analytics Science And Technology, 2006 IEEE Symposium On*, pages 199–206, Oct. 2006.

[12] John W. Tukey. *Exploratory data analysis*. 1977.

[13] L. Wilkinson, A. Anand, and R. Grossman. Graph-theoretic scagnostics. In *Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on*, pages 157–164, 23-25 Oct. 2005.

Reviewers' Index

I
Pedro Isaias

J
Hewijin Christine Jiau
Qun Jin
Joemon Jose
Erland Jungert

K
Wolfgang Klas
Lars Knipping
Hideki Koike
Jun Kong
Zenon Kulpa
Yau-Hwang Kuo

L
Robert Laurini
Jen Juan Li
Fuhua Lin
Alan Liu
Chien-Tsai Liu
Chung-Fan Liu
Jonathan Liu
Benjamin Lok

M.
Andrea Marcante
Paolo Maresca
Kim Marriott
Sergio Di Martino
Rym Mili
Andrea Molinari
Piero Mussio

N
Paolo Nesi

O
Vincent Oria

P
Sethuraman Panchanathan
Luca Paolino
Joseph J. Pfeiffer
Antonio Piccinno
Di Bitonto Pierpaolo
Sofie Pilemalm
Fabio Pittarello
Valentin Plenk
Beryl Plimmer
Giuseppe Polese

R
Syed M. Rahman
Steven P. Reiss
Teresa Roselli

S
Nicoletta Sala
Giuseppe Santucci
Giuseppe Scanniello
Monica Sebillo
Timothy K. Shih
Peter Stanchev
Gem Stapleton
David Stotts
Hui-Kai Su
Yu-Huei Su
Nik Swoboda

T
Genny Tortora

U
Joseph E. Urban

V

Athena Vakali
Athanasios Vasilakos
Rossano Veronica
Giuliana Vitiello
Son T. Vuong

W

Ellen Walker
KaiYu Wan
Mingzoo Wu

Y

Kazuo Yana
Chi-Lu Yang
Atsuo Yoshitaka

Z

Kang Zhang

Authors' Index

G

H

I

J

K

L

M

N

O

P

R

S

T

V

W

Y

Z

DMS 2010 Call For Papers

The Sixteenth International Conference on Distributed Multimedia Systems

September 10 - September 12, 2010
Organized by
Knowledge Systems Institute Graduate School

SCOPE
The DMS conference is an international conference series, which covers a wide spectrum of technique discussions, demonstrations, and student program/paper contests in the fields of distributed multimedia computing. Started in 1994, the series of conference has been held at Hawaii, Hong Kong, Vancouver, Taipei, Aizu-Wakamatsu, Japan, San Francisco, Miami and Canada.

TOPICS
Distributed Multimedia Technology
- media coding and compression standards
- media streaming and QoS control
- digital rights management
- middleware and OS support
- privacy and security issues
- mobile devices and wireless networks
- sensor networks
- context sensing and processing

Distributed Multimedia Models and Systems
- human-computer interaction issues
- situated computing
- languages for distributed multimedia: Java, XML, X3D
- software engineering issues
- web services and multi-agent systems
- multimedia databases and information systems
- multimedia indexing and retrieval systems
- distributed multimedia systems evaluation
- multimedia authoring

Applications of Distributed Multimedia Systems
- distributed collaboration
- cultural heritage preservation, management and fruition
- distance and lifelong learning
- emergency and safety management
- environment control and management
- e-commerce and e-government applications
- health care management and disability assistance
- intelligent multimedia computing
- social networking
- virtual, mixed and augmented reality

The use of prototypes and demonstration video for presentations is encouraged.

WORKSHOPS AND SPECIAL SESSIONS
This year, DMS 2010 will be held in conjunction with workshops, conferences and special sections. Papers submitted to workshops or special sessions are invited by the program committee chairs of the workshops/sections. This joint organization aims to collect research results from different perspectives. The following workshops and conferences are being planned:
International Workshop on Visual Languages and Computing
International Workshop on Mobile Systems, E-Commerce and Agent Technology
International Workshop on Distance Education Technology

Please contact the conference program co-chairs if you are interested in organizing a workshop or a special session. A one-page proposal with the following items is required:
Title of the workshop/special session
Name of the program committee chair(s)
A list of program committee members
E-mail address of the corresponding program committee chair
A brief description of the theme

Each special session should include at least 5 invited (or accepted) papers. Each workshop should include at least 2 special sessions (10 papers). Paper review process of the workshop/special session will be arranged by the individual program committee chair. It is the responsibility of the individual chairs to ensure the quality of papers in the workshop/special sessions. The schedule of workshops/special sessions will be the same as the main conference (See Important Dates below). Each individual program committee chair can set up the Web site of the workshop/special session. However papers will be printed in the same volume as the main conference.

SPECIAL ISSUE OF JOURNALS
The DMS conference is closely coordinated with the International Journal of Software Engineering and Knowledge Engineering, the Journal of Visual Languages and Computing and the International Journal of Distance Education Technologies. Usually the best ranked papers from the conference, after rigorous reviews, extensive revisions and further enhancements, will appear in one or more special issues of the above journals. Papers suggested for the special issue(s) will be reviewed by external reviewers following the standard procedure of review stipulated by the respective journal.

INFORMATION FOR AUTHORS
Papers must be written in English. An electronic version (Postscript, PDF, or MS Word format) of the full paper should be submitted using the following URL: http://conf.ksi.edu/dms10/submit/SubmitPaper.php. Please use Internet Explorer as the browser. Manuscript must include a 200-word abstract and no more than 6 pages of IEEE double column text (include figures and references).

INFORMATION FOR REVIEWERS
Papers submitted to DMS'10 will be reviewed electronically. The users (webmaster, program chair, reviewers...) can login using the following URL:
http://conf.ksi.edu/dms10/review/pass.php.
If you have any questions or run into problems, please send e-mail to: dms10@ksi.edu.

DMS 2010 Conference Secretariat
Knowledge Systems Institute Graduate School
3420 Main Street
Skokie, IL 60076 USA
Tel: 847-679-3135
Fax: 847-679-3166
E-mail: dms10@ksi.edu

If you cannot submit electronically, please send four copies of the complete manuscript to the above postal address.

IMPORTANT DATES
April 30, 2010 *Paper submission due*
May 31, 2010 *Notification of acceptance*
June 15, 2010 *Final manuscript due*
June 15, 2010 *Early conference registration due*

Proceedings of the Fifteenth International Conference on Distributed Multimedia Systems

DMS 2009

Printed by
Knowledge Systems Institute
3420 Main Street
Skokie, Illinois 60076
(847) 679-3135
info@ksi.edu
www.ksi.edu
Printed in USA, 2009
ISBN 1-891706-25-X (paper)

Co-Sponsored by:

Eco Controllo SpA, Italy
University of Salerno, Italy
University Ca' Foscari in Venice, Italy
Knowledge Systems Institute Graduate School, USA