DMS 2016

Salerno, Italy
November 25-26

Proceedings of the Twenty-Second International Conference on Distributed Multimedia Systems

and

Journal of Visual Languages and Sentient Systems, Volume 2, 2016

PROCEEDINGS

DMS 2016

The 22nd International Conference on Distributed Multimedia Systems

Sponsored by

KSI Research Inc., Knowledge Systems Institute Graduate School and Datonix

Technical Program

November 25 – November 26, 2016

Grand Hotel Salerno, Italy

Organized by

KSI Research Inc. and Knowledge Systems Institute Graduate School, USA

Copyright ⓒ 2016 by KSI Research Inc. and Knowledge Systems Institute Graduate School

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written consent of the publisher.

ISBN: 1-891706-40-3

ISSN: 2326-3261 (print)

2326-3318 (online)

DOI: 10.18293/DMS2016

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: dms@ksiresearch.org

Web: http://www.ksi.edu

Proceedings preparation, editing and printing are sponsored by KSI Research Inc. and Knowledge Systems Institute Graduate School, USA.

Printed by KSI Research Inc. and Knowledge Systems Institute Graduate School, USA.

FOREWORD

Welcome to the 22nd International Conference on Distributed Multimedia Systems (DMS 2016) that takes place this year in Grand Hotel Salerno, Italy. Researchers from academia and industry from around the world meet to discuss issues, ideas, and innovations within the areas of multimedia and sentient systems.

The conference is organized in sessions which focus on many specialized topics. Exchange of ideas, discussions, research results, and experiences in the longstanding history of the conference have had a positive influence on the research in the past, and we believe that they will have a positive influence this year as well, thanks to the quality of the meeting and the research contributions from researchers from many countries. With the high quality of this year's technical program, the DMS community will continue to be an important venue and a source of new ideas and innovations. However the DMS conference is not limited simply to research on multimedia sentient systems, and with the support of two additional workshops, one on Distance Education Technology (DET 2016) and one on Visual Languages and Computing (VLC 2016), the DMS conference continues to provide an international forum for discussion that expands also into the areas of education and visual languages.

We have received 33 submissions and the paper selection was based upon a rigorous review process, with an acceptance rate for full papers of 48%. We are expecting authors and guests from 8 countries: Algeria, China, France, Germany, Italy, New Zealand, UK, and USA. This year's conference program contains contributions of high quality research papers and short papers to discuss ongoing research activities and applications. Two timely keynotes from Gianluigi Riccio and Ferdinando Gorga, both reflecting the views from the industry practitioners, are also included in the technical program.

Starting last year, The DMS proceedings also contain the Journal of Visual Languages and Sentient Systems (JVLSS). Volume I of JVLSS was published together with DMS2015 Proceedings, and Volume II of JVLSS is published together with DMS2016 Proceedings. In addition to this, up to 9 papers will be invited and further reviewed for possible inclusion in the special issues on best papers from DMS2016, to be published in December 2016 in Journal of Visual Languages and Computing (JVLC). Only papers presented at the DMS2016 will be considered for this special issue. Invitation will be made after the DMS2016 conference.

As Program Co-Chairs, we would like to express all our gratitude and appreciation to the Steering Committee Chair Dr. S.K. Chang for his support, dedication to the conference, and his invaluable experience. However the high quality of DMS 2016 technical program would not have been possible without the tireless efforts of many individuals. First of all, we would like to thank the entire Steering Committee for their continuous support and guidance; the entire Program Committee whose invaluable, attentive, and timely work has made possible the creation of a high quality technical program. Then, we would like to extend our sincere appreciation to all the authors who have submitted their papers to the conference, thus contributing with their work and ideas to the success of this venue. Last but not least, we like to acknowledge the important contribution of the KSI Research staff whose assistance and support has been truly remarkable throughout the entire organization process.

This will also be the last year that the conference bears the name "Distributed Multimedia Systems". Starting from next year, the conference will be renamed "Distributed Multimedia Systems, Visual Languages and Sentient Systems" (DMSVLSS) to better position the conference. The two workshops VLC and DET will be combined with the main conference. Finally, the conference will be co-located with the International Conference on Software Engineering and Knowledge Engineering (SEKE).

On behalf of the Program Committee, we are delighted to extend to you our warm welcome to the 22nd International Conference on Distributed Multimedia Systems (DMS 2016). We hope that you will find this year's conference an exciting place for exchanging ideas, for fostering new projects, and a rewarding place for your research challenges. We wish you a nice staying in Salerno and we hope that you will find some time to enjoy, among other things, the beauty that the city and the country have to offer.

Giuseppe Polese and Vincenzo Deufemia
DMS 2016 Program Co-Chairs

DMS 2016

The 22nd International Conference on Distributed Multimedia Systems

November 25 – November 26, 2016

Grand Hotel Salerno, Salerno, Italy

Conference Organization

DMS'16 Conference Co-Chairs

Angela Guercio, Kent State University, USA; Conference Co-Chair
Mahbubur Syed, Minnesota State University, USA; Conference Co-Chair

DMS'16 Honorary Chair

Genoveffa Tortora, University of Salerno, Italy; Honorary Chair

DMS'16 Steering Committee Chair

Shi-Kuo Chang, University of Pittsburgh, USA; Steering Committee Chair

DMS'16 Steering Committee

Paolo Nesi, University of Florence, Italy; Steering Committee Member
Kia Ng, University of Leeds, UK; Steering Committee Member

DMS'16 Program Co-Chairs

Giuseppe Polese, University of Salerno, Italy; Program Co-Chair
Vincenzo Deufemia, University of Salerno, Italy; Program Co-Chair

DMS'16 Program Committee

Arvind K. Bansal, Kent State University, USA
Andrew Blake, University of Brighton, UK

Ing-Ray Chen, Virginia Tech (VPI&SU), USA
Shu-Ching Chen, Florida International University, USA
Loredana Caruccio, University of Salerno, Italy
William Cheng-Chung Chu, Tunghai University, Taiwan
F. Colace, University of Salerno, Italy
Gennaro Costagliola, Univ of Salerno, Italy
Alfredo Cuzzocrea, ICAR-CNR and University of Calabria, Italy
Andrea De Lucia, Univ. of Salerno, Italy
Tiansi Dong, Bonn-Aachen International Center for Information Technology, Germany
David H. C. Du, Univ. of Minnesota, USA
Larbi Esmahi, National Research Council of Canada, Canada
Daniela Fogli, Universita degli Studi di Brescia, Italy
Kaori Fujinami, Tokyo University of Agriculture and Technology, Japan
David Fuschi, Brunel University, UK
Ombretta Gaggi, Univ. of Padova, Italy
Nikolaos Gkalelis, Informatics & Telematic Institute, Greece
Angela Guercio, Kent State University, USA
Carlos A. Iglesias, Intelligent Systems Group, Spain
Erland Jungert, Linkoping University, Sweden
Yau-Hwang Kuo, National Cheng Kung University, Taiwan
Fuhua Lin, Athabasca University, Canada
Alan Liu, National Chung Cheng Univeristy, Taiwan
Jonathan Liu, University of Florida, USA
Max North, Southern Polytechnic State University , USA
Sethuraman Panchanathan, Arizona State Univ., USA
Antonio Piccinno, Univ. of Bari, Italy
Giuseppe Polese, University of Salerno, Italy
Buntarou Shizuki, University of Tsukuba, Japan
Peter Stanchev, Kettering University, USA
Genny Tortora, University of Salerno, Italy
Atsuo Yoshitaka, JAIST, Japan
Ing Tomas Zeman, Czech Technical University, Czech Republic
Kang Zhang, University of Texas at Dallas, USA

DET'16 Workshop Co-Chairs

Maiga Chang, Athabasca University, Canada; DET'16 Workshop Co-Chair
Rita Francese, University of Salerno, Italy; DET'16 Workshop Co-Chair

DET'16 Program Chair

Michele Risi, University of Salerno, Italy; DET'16 Program Chair

DET'16 Program Committee

Tim Arndt, Cleveland State University, USA
Maiga Chang, Athabasca University, Canada
Yuan-Sun Chu, National Chung Cheng University, Taiwan
Mauro Coccoli, University of Genova, Italy
Luigi Colazzo, University of Trento, Italy
Rita Francese, University of Salerno, Italy
Angelo Gargantini, University of Bergamo, Italy
Angela Guercio, Kent State University, USA
Robert Heller, Athabasca University, Canada
Pedro Isaias, Open University, Portugal
Paolo Maresca, University Federico II, Napoli, Italy
Andrea Molinari, University of Trento, Trento, Italy
Mario Arrigoni Neri, University of Bergamo, Italy
Ignazio Passero, University of Salerno, Italy
Michele Risi, University of Salerno, Italy
Teresa Roselli, University of Bari, Italy
Veronica Rossano, University of Bari, Italy
Giuseppe Scanniello, University of Basilicata, Italy
Lidia Stanganelli, Freelance Software Engineer, Italy

VLC'16 Workshop Chair

Gem Stapleton, University of Brighton, UK; VLC'16 Workshop chair

VLC'16 Program Chair

Peter Chapman, University of Brighton, UK; Program Chair

VLC'16 Program Committee

Bilal Alsallakh, Vienna University of Technology, Austria
Danilo Avola, University of Rome, Italy
Paolo Bottoni, Universita Sapienza, Italy
Paolo Buono, University of Bari, Italy
Alfonso F. Cardenas, University of California, USA
Peter Chapman, University of Brighton, UK
Gennaro Costagliola, University of Salerno, Italy
Aidan Delaney, University of Brighton, UK
Vincenzo Deufemia, University of Salerno, Italy

Filomena Ferrucci, University of Salerno, Italy
Andrew Fish, University of Brighton, UK
Vittorio Fuccella, University of Salerno, Italy
Joaquim A Jorge, Instituto Suoerior Tecnico, Portugal
Levent Burak Kara, Carnegie Mellon University, USA
Robert Laurini, University of Lyon, France
Jennifer Leopold, Missouri University of Science & Technology, USA
Luana Micallef, Helsinki Institute for Information Technology, Finland
Nikolay Mirenkov, University of Aizu, Japan
Joseph J. Pfeiffer, Jr., New Mexico State University, USA
Peter Rodgers, University of Kent, UK
Monica Sebillo, University of Salerno, Italy
Gem Stapleton, University of Brighton, UK
Franklyn Turbak, Wellesley College, USA
Giuliana Vitiello, University of Salerno, Italy

DMS/DET/VLC'16 Publicity Co-Chairs

Kao-Shing Hwang, National Chung Cheng University, Taiwan; Publicity Co-Chair
Lidia Stanganelli, University of Napoli Federico II, Italy; Publicity Co-Chair

Keynote 1

Is Data Preparation the next Big Data disruption?

Gianluigi Riccio
CEO, Datonix Spa, Italy
E-mail: gianluigi@datonix.it

Abstract: Data Preparation is growing essentially because the Data Scientists community is expanding. Data Preparation technologies have finally been recognized from the IT industry analyst community. In 2015, market resulted 500M USD size coming from a list of 20 to 30 vendors. In addition, recently Gartner forecasted Dprep Market should reach 1B USD size by 2018. Today, Enterprises need pragmatic transformation strategies to advance goals and recognize emerging opportunities. To meet transformation opportunities, companies are introducing the data driven culture, modernizing their core infrastructure and investing wisely in new capabilities. As a result it seems that every business might be an analytics business and every user might have access to analytics to succeed. The presentation will explore following critical issues:

- how companies are using self-service data preparation,
- how they will implement next generation data discovery techniques on data and multimedia information
- how modern data science is evolving to deliver more agile and high value analytics to all users on multimedia information.

About the Speaker: Gianluigi Riccio is founder and chief technology officer @ datonix.it He is a noted speaker and author on Global 5000 enterprise IT issues and has been one of the most quoted industry analyst on the topics of data base technology. Prior to founding datonix, he was founder of a quick growing data base consulting firm in southern Italy. During the nineteen nineties, he was research director of META Group's largest research advisory practice for 8 years. He has also held roles in leading vendor and user organizations, including executive and managerial positions at QueryObject Corp., Aeritalia Saipa, Olivetti OPE, Gartner Group, and Eureka Consortium. Gianluigi received his electronic engineering degree from University of Naples.

Keynote 2

The universities as greatness generators, the companies as technology generators - how the technological innovation affects the future of students and companies

Ferdinando Gorga
IBM Bluemix Cloud Architect
University Ambassador
IBM Italy, Roma
ferdinando_gorga@it.ibm.com

Abstract: In blocks programming languages (such as Scratch, Blockly, App Inventor, Snap!, Pencil Code, Alice/Looking Glass, AgentSheets/AgentCubes), programs are constructed by connecting visual blocks shaped like puzzle pieces. Through activities like Code.org's Hour of Code and both online and traditional courses, these languages have become extremely popular ways to introduce programming and computational thinking to tens of millions of people of all ages and backgrounds. By lowering barriers to programming in key programming language dimensions (syntax, static semantics, and dynamic semantics), blocks languages are helping to democratize programming by putting the power of programming in the hands of nonexperts. In my talk, I will focus on blocks language work done in the context of MIT App Inventor and the Wellesley College TinkerBlocks research project. Despite recent advances in blocks languages, there are still many challenges to address, including enhancing their usability and expressiveness, developing paths for transitioning to more traditional programming, and dealing with the perception that they are just toy languages for kids. I encourage members of the DMS community to join me in investigating these challenges.

About the Speaker: Graduated in Computer Science in Salerno University, specialized in Software engineering and programming languages, Ferdinando Gorga is an IBM evangelist. His job is to understand the needs of companies and to find solutions to improve and protect their business. At present he is a "Bluemix Cloud Architect" for the IBM Cloud Business Unit and he is also an University Ambassador for the IBM University Academic Initiative. Ferdinando lives from Rome and Cilento, in Italy, His main interests are: his work, science, photography, travels, motorbikes, sea.

Table of Contents

Foreword iii

Conference Organization v

KeyNote 1: Is Data Preparation the next Big Data disruption?
Mr Gianluigi Riccio ix

KeyNote 2: The universities as greatness generators, the companies as technology generators - how the technological innovation affects the future of students and companies
Mr Ferdinando Gorga x

VLSS

Recommender Systems and Social Networks: an application in Cultural Heritage 1
Flora Amato, Vincenzo Moscato, Antonio Picariello and Giancarlo Sperlì

VisCFSM: Visual, Constraint-Based, Frequent Subgraph Mining 9
Jennifer Leopold, Nathan Eloe and Chaman Sabharwal

A toolkit for knot diagram sketching, encoding and re-generation 16
Gennaro Costagliola, Mattia De Rosa, Andrew Fish, Vittorio Fuccella, Rafiq Saleh and Sarah Swartwood

Visualizing student engagement in e-learning environment 26
Enrica Pesare, Teresa Roselli and Veronica Rossano

Is e-learning ready for big data? And how big data would be useful to e-learning ? 34
Paolo Maresca and Andrea Molinari

Supporting Mobile Development Project-Based learning by Software Project and Product Measures 41
Rita Francese, Carmine Gravino, Michele Risi, Giuseppe Scanniello and Genoveffa Tortora

Towards Formal Multimodal Analysis of Emotions for Affective Computing 48
Mehdi Ghayoumi, Arvind Bansal and Maha Thafar

Parameter Calibration Method of Microscopic Traffic Flow Simulation Models based on Orthogonal Genetic Algorithm 55
Yanfang Yang, Yong Qin, Honghui Dong and Qing Zhang

DMS-1

A Novel Priority-based Deadlock Detection and Resolution Algorithm in Mobile Agent Systems 61
Wei Lu, Yong Yang, Weiwei Xing, Liqiang Wang and Xiaoping Che

An Entropy based Product Ranking Algorithm using Reviews and Q&A Data 69
Bushra Anjum and Chaman Sabharwal

Novel Group Detection and Analysis Method Based on Automatic and Fast Density Clustering (S) 77
Weiwei Xing, Ke Jin and Peng Bao

Towards An Effective and Efficient Approximation Algorithm for Advanced Computer Vision Applications based on Two-Dimensional Dynamic Programming (S) 84
Alfredo Cuzzocrea, Enzo Mumolo, Giorgio Mario Grasso and Gianni Vercelli

VLC-1

VisCFSM: Visual, Constraint-Based, Frequent Subgraph Mining 9
Jennifer Leopold, Nathan Eloe and Chaman Sabharwal

An Introduction to Geographic Rules (S) 91
Robert Laurini, Sylvie Servigne and Franck Favetta

Automatic Generation of an Interactive, Real time, Web-based Maps from Sensor-based Geographic Databases (S) 98
Zina Bouattou, Robert Laurini and Hafida Belbachir

VLC-2

A toolkit for knot diagram sketching, encoding and re-generation 16
Gennaro Costagliola, Mattia De Rosa, Andrew Fish, Vittorio Fuccella, Rafiq Saleh and Sarah Swartwood

VeCVL: A Visual Language for Version Control (S) 105
Nathan Eloe, Denise Case and Jennifer Leopold

A Browser-based IDE for the MUzECS Platform (S) 112
Omokolade Hunpatin, Casey O'Hare, Ryan Thomas and Dennis Brylow

DMS-2

Social Network Models for the TDR System 119
Shikuo Chang

Distributing the User Interface Logics along Actionable Components: the EFESTO Approach 129
Giuseppe Desolda, Carmelo Ardito, Maristella Matera and Maria Francesca Costabile

Efficient Unsupervised Behavioral Segmentation of Human Motion Capture Data (S) 138
Xiaomin Yu, Weibin Liu and Weiwei Xing

Towards a Dialogue System Supporting Automatic Event Identification (S) 148
Yibin Jiang, Tiansi Dong, Armin B. Cremers and Joachim Köhler

DET-1

Is e-learning ready for big data? And how big data would be useful to e-learning ? 34
Paolo Maresca and Andrea Molinari

A Tool for the Semantic Analysis and Recommendation of videos in e-learning (S) 153
Mauro Coccoli and Gianni Vercelli

Teaching Computer Programming Through Hands-on Labs on Cognitive Computing (S) . . 158
Mauro Coccoli, Paolo Maresca and Lidia Stanganelli

DMS-3

Recommender Systems and Social Networks: an application in Cultural Heritage 1
Flora Amato, Vincenzo Moscato, Antonio Picariello and Giancarlo Sperlì

A Context-Aware Approach for a Collaborative, Pervasive and Adaptive Digital Storytelling 165
Francesco Colace, Mario Casillo, Saverio Lemma, Marco Lombardi and Massimo De Santo

DET-2

Visualizing student engagement in e-learning environment 26
Enrica Pesare, Teresa Roselli and Veronica Rossano

Scaffolding Version Control into the Computer Science Curriculum 175
Denise M. Case, Nathan W. Eloe and Jennifer L. Leopold

Supporting Mobile Development Project-Based learning by Software Project and Product Measures 41
Rita Francese, Carmine Gravino, Michele Risi, Giuseppe Scanniello and Genoveffa Tortora

DMS-4

Parameter Calibration Method of Microscopic Traffic Flow Simulation Models based on Orthogonal Genetic Algorithm 55
Yanfang Yang, Yong Qin, Honghui Dong and Qing Zhang

Research on Performance Degradation Assessment Method of Train Rolling Bearings under incomplete data (S) 184
Xuejun Zhao, Yong Qin, Dandan Wang, Zhipeng Wang and Limin Jia

Temporal-Based Risk Forecasting Approach for Key Areas on Surveillance Sensor Networks of High-Speed Railway Transport Hub (S) 190
Zhengyu Xie, Yong Qin and Li Wang

DMS-5

Smart Home Control through Unwitting Trigger-Action Programming 194
Daniela Fogli, Matteo Peroni and Claudia Stefini

AP Positioning for Estimating People Flow as Origin Destination Matrix for Smart Cities 202
Paolo Nesi, Pierfrancesco Bellini and Daniele Cenni

Towards Formal Multimodal Analysis of Emotions for Affective Computing 48
Mehdi Ghayoumi, Arvind Bansal and Maha Thafar

Visualizing Transportation Routes for Data Analysis in Logistics (S) 210
Paolo Buono

Authors' Index A-1
Progam Committee Reviewers' Index A-4
External Reviewers' Index A-7

Note:
(S) indicates a short paper.

Journal of

Visual Languages and Sentient Systems

Volume 2, 2016

Journal of
Visual Languages and Sentient Systems

Editor-in-Chief

Shi-Kuo Chang, University of Pittsburgh, USA

Co-Editors-in-Chief

Gennaro Costagliola, University of Salerno, Italy

Paolo Nesi, University of Florence, Italy

Gem Stapleton, University of Brighton, UK

Franklyn Turbak, Wellesley College, USA

An Open Access Journal published by

KSI Research Inc.

and

Knowledge Systems Institute Graduate School

VLSS Editorial Board

Tim Arndt, Cleveland State University, USA

Alan F. Blackwell, University of Cambridge, United Kingdom

Paolo Bottoni, University of Rome, Italy

Francesco Colace, Univeristy of Salerno, Italy

Maria Francesca Costabile, University of Bari, Italy

Philip T. Cox, Dalhousie University, Canada

Martin Erwig, Oregon State University, USA

Vittorio Fuccella, University of Salerno, Italy

Angela Guercio, Kent State University, USA

Erland Jungert, Swedish Defence Research Establishment, Sweden

Kamen Kanev, Shizuoka University, Japan

Robert Laurini, University of Lyon, France

Jennifer Leopold, Missouri University of Science & Technology, USA

Mark Minas, University of Munich, Germany

Brad A. Myers, Carnegie Mellon University, USA

Joseph J. Pfeiffer, Jr., New Mexico State University, USA

Genny Tortora, University of Salerno, Italy

Kang Zhang, University of Texas at Dallas, USA

Copyright ⓒ 2016 by KSI Research Inc. and Knowledge Systems Institute Graduate School

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written consent of the publisher.

ISBN: 1-891706-40-3

ISSN: 2326-3261 (print)

2326-3318 (online)

DOI: 10.18293/DMS2016

Proceedings preparation, editing and printing are sponsored by KSI Research Inc. and Knowledge Systems Institute Graduate School, USA.

Printed by KSI Research Inc. and Knowledge Systems Institute Graduate School, USA.

Journal of
Visual Languages and Sentient Systems

Volume 2, 2016

Table of Content

Information for Authors IV

Regular Papers

Flora Amato, Vincenzo Moscato, Antonio Picariello and Giancarlo Sperlì, "Recommender Systems and Social Networks: An Application in Cultural Heritage" 1

Jennifer Leopold, Nathan Eloe and Chaman Sabharwal, "VisCFSM: Visual, Constraint-Based, Frequent Subgraph Mining" 9

Gennaro Costagliola, Mattia De Rosa, Andrew Fish, Vittorio Fuccella, Rafiq Saleh and Sarah Swartwood, "A Toolkit for Knot Diagram Sketching, Encoding and Re-generation" 16

Enrica Pesare, Teresa Roselli and Veronica Rossano, "Visualizing Student Engagement in e-learning Environment" 26

Paolo Maresca and Andrea Molinari, "Is e-learning Ready for Big Data? And How Big Data Would Be Useful to e-learning ?" 34

Rita Francese, Carmine Gravino, Michele Risi, Giuseppe Scanniello and Genoveffa Tortora, "Supporting Mobile Development Project-Based Learning by Software Project and Product Measures" 41

Mehdi Ghayoumi, Arvind Bansal and Maha Thafar, "Towards Formal Multimodal Analysis of Emotions for Affective Computing" 48

Yanfang Yang and Yong Qin, "Parameter Calibration Method of Microscopic Traffic Flow Simulation Models based on Orthogonal Genetic Algorithm" 55

INFORMATION FOR AUTHORS

The Journal of Visual Languages and Sentient Systems (VLSS) is intended to be a forum for researchers, practitioners and developers to exchange ideas and research results, for the advancement of visual languages and sentient multimedia systems. Sentient systems are distributed systems capable of actively interacting with the environment by gathering, processing, interpreting, storing and retrieving multimedia information originated from sensors, robots, actuators, websites and other information sources. In order for sentient systems to function efficiently and effectively, visual languages may play an important role.

VLSS publishes research papers, state-of-the-art surveys, review articles, in the areas of visual languages, sentient multimedia systems, distributed multimedia systems, sensor networks, multimedia interfaces, visual communication, multi-media communications, cognitive aspects of sensor-based systems, and parallel/distributed/neural computing & representations for multimedia information processing. Papers are also welcome to report on actual use, experience, transferred technologies in sentient multimedia systems and applications. Timely research notes, viewpoint articles, book reviews and tool reviews, not to exceed three pages, can also be submitted to VLSS.

Manuscripts shall be submitted electronically to VLSS. Original papers only will be considered. Manuscripts should follow the double-column format and be submitted in the form of a pdf file. Page 1 should contain the article title, author(s), and affiliation(s); the name and complete mailing address of the person to whom correspondence should be sent, and a short abstract (100-150 words). Any footnotes to the title (indicated by *, +, etc.) should be placed at the bottom of page 1.

Manuscripts are accepted for review with the understanding that the same work has not been and will not be nor is presently submitted elsewhere, and that its submission for publication has been approved by all of the authors and by the institution where the work was carried out; further, that any person cited as a course of personal communications has approved such citation. Written authorization may be required at the Editor's discretion. Articles and any other material published in VLSS represent the opinions of the author(s) and should not be construed to reflect the opinions of the Editor(s) and the Publisher.

Paper submission should be through https://www.easychair.org/conferences/?conf=dmsvlss2017

For further information contact: vlss@ksiresearch.org

Recommender Systems and Social Networks: an application in Cultural Heritage

Flora Amanto, Vincenzo Moscato, Antonio Picariello, Giancarlo Sperlí
Dip. di Ingegneria Elettrica e Tecnologie dell'Informazione
University of Naples "Federico II", Naples, Italy
{flora.amato, vmoscato, picus, giancarlo.sperli}@unina.it

Abstract

In the last decade Recommender Systems *have become useful tools helping users to find "what they need" from considerable amount of data. One of the more obvious applications of such systems in the Cultural Heritage domain is to assist users when visiting cultural environments (such as museums, archaeological sites, old town centers and so on), providing a multimedia guide that is able to dynamically suggest relevant information available in multiple web repositories (e.g. multimedia sharing systems and on-line social networks). In this paper, we propose a novel recommendation approach that combines several aspects of users - i.e. their* preferences *(usually in the shape of items' metadata) and* interactions *within a social community modeled using hypergraphs - together with items'* multimedia features *and* context information *within a general framework that can support different applications (touristic guiding services for museums, visiting paths recommendation for old town centers and archeological sites, etc.). Preliminary experiments on user satisfaction show how our approach provides very promising and interesting results.*

1 Introduction

The development and promotion of worldwide Cultural Heritage using Information and Communication Technologies (ICT) represent nowadays an important research issue with a variety of potential applications.

In the last decade, such technologies have radically changed the purpose of Cultural Heritage exhibitions that is rapidly moving from an old vision, providing a tourist with static information consisting of a large amount of cultural signs, to novel *personalized services*, matching the visitors' personal goals and behaviors by considering their cultural needs and preferences and context information.

Indeed, users' experience could be surely enhanced if, instead of using classic "tourist" devices, they could be embedded in a cultural environment with a number of functionalities for representing the relevant information derived from the available digital sources, such as text descriptions, pictures, and videos. In this way, tourists would be given the opportunity of enjoying multimedia stories in real time, thus enriching their cultural knowledge.

From the other hand, we are assisting to an explosive and amazing increase of digital information, and as a consequence, more and more huge data collections of different nature are widely available and have constrained users necessarily to deal with this ocean of information to find "what they need". In particular, on-line social networks (e.g. Facebook) and multimedia sharing systems (e.g. YouTube, Flickr, Panoramio, Instagram, etc.), together with open digital libraries and archives (e.g. DBpedia), constitute the main multimedia information sources that can be considered "useful" for tourists when they visiting cultural environments such as museums, archaeological sites, old town centers and so on.

As well known, *Recommender Systems* have been introduced to facilitate the browsing of such collections, thus realizing the transition in the Web from the *search* to the *discovery* paradigm.

Generally, recommender systems help people in retrieving information that match their preferences by recommending products or services from a large number of candidates, and support people in making decisions in various contexts: what items to buy, which movie to watch, which music to listen, what travels to do, or even who they can invite to their social network, just to make some examples [24, 25].

One of the more obvious applications of such systems in the Cultural Heritage domain is to assist users

DOI reference number: 10.18293/DMS2016-001

when visiting cultural environments, providing a *multimedia guide* that is able to dynamically suggest relevant information available in multiple web repositories.

Formally, a recommender system deals with a set of *users* $U = \{u_1 \ldots, u_m\}$ and a set of *items* $O = \{o_1, \ldots, o_n\}$. For each pair (u_i, o_j), a recommender can compute a *score* (or a *rank*) $r_{i,j}$ that measures the expected interest of user u_i in item o_j (or the expected utility of item o_j for user u_i), using a *knowledge base* and a *ranking* algorithm that generally could consider different combinations of the following characteristics: (i) user preferences and past behavior, (i) preferences and behavior of the user community, (iii) items' features and how they can match user preferences, (iv) user feedbacks, (v) context information (i.e. user location, observed items, weather and environmental conditions, etc.) and how recommendations can change together with the context.

In the literature, surveys on recommender systems usually classify the different kinds of approaches in four main categories: *content-based* [22, 29, 30] (with their extensions to deal multimedia data and their features [19, 13, 20]) , *collaborative filtering* [2, 29, 23, 17] (with their customizations to take into account social elements as user reviews and opinions [32, 28, 18, 27, 11, 21, 9]), *hybrid* [26] and *context aware* [10, 15, 16] techniques. Finally, a recent category of recommenders, named *Large Scale Recommender Systems* (LSRS) [31], calls for new capabilities of such applications to deal with very large amount of data with respect to scalability and efficiency issues.

In our opinion, modern recommending applications have to take into account in some way all the above characteristics to provide useful and reliable recommendations both for virtual and physical environments. To this goal, the last generation of recommender systems is usually composed by one or more of the following components [25].

A *pre-filtering* module that selects for each user u_i a subset $O_i^c \subset O$ containing items that are good candidates to be recommended; such items usually match user preferences and needs.

A *ranking* module that assigns w.r.t. user u_i a rank $r_{i,j}$ to each candidate item o_j in O_i^c using the well-known recommendation techniques (i.e., *content-based*, *collaborative filtering* and *hybrid* approaches) that can exploit in several ways items' features and users' preferences, feedbacks (in the majority of cases in terms of *ratings*) and behavior.

A *post-filtering* module that dynamically excluds, for each user u_i, some items from the recommendations' list; in this way, a new set $O_i^f \subseteq O_i^c$ is obtained on the base of user feedbacks, other contextual information (such as data coming from the interactions between the user and the application) and possible additional constraints.

In this paper, we propose a novel recommendation approach that combines several aspects of users - i.e. their *preferences* (in the shape of items' metadata) and *interactions* (user to user and user to content) within a social community modeled using hypergraphs - together with items' *multimedia features* and *context information* within a general framework that can support different applications (touristic guiding services for museums, visiting paths recommendation for old town centers and archeological sites, etc.).

In other words, it is the user with his/her preferences (in the pre-filtering stage) and actions (in the post-filtering stage) to drive the recommendation process towards the real useful items among those that a social community considers the "best ones" (computed in the ranking stage), as in a collaborative filtering approach, where a user "learns by the others" the item utility, on the base of an *influence* measure.

The paper is organized as follows. Section 2 provides a functional overview of our system and describes the proposed strategy for recommendation. Section 3 illustrates a system customization for a tourist multimedia guide, reporting some implementation details. Section 4 reports preliminary experimental results, and provides a comparison with other recommendation techniques. Finally, Section 5 gives some concluding remarks and discusses future work.

2 The framework

2.1 System Overview

Figure 1 describes at a glance an overview of the proposed system.

Multimedia data to be recommended are retrieved by a *Wrapper* component that is composed by several modules. The *Crawler* is responsible of: (i) periodically accessing to the items' repositories (e.g., *Instagram*, *Flickr*, *Panoramio*, *Google Images*, *YouTube*,*Facebook*, *DBpedia* etc.), (ii) extracting for each item all the *features* (e.g., metadata, multimedia descriptions, etc.) and other information (e.g. user preferences, comments, time-stamped items' observations and all the different interactions between users and objects). A part of such information will be then exploited by the *Hypergraph Learning* module to build the *hypergraph* modeling the entire *Multimedia Social Network* (MSN)[6]. After the wrapping phase, all the information are stored in the *Knowledge Base* of the

2

system. In particular, it is composed by: (i) the *Multimedia Social Network Hypergraph*, (ii) the *Items DB* containing items with all the related features, (iii) *User Profiles* containing user preferences, (iv) *Contextual Data* containing some additional context information (e.g. user location, weather conditions, etc.).

Multimedia items are then grouped by the related *cultural Points Of Interest* (POIs): e.g. paintings of museum rooms, buildings in ancient ruins or in an old town center, etc.

The *Recommender Engine* provides a set of recommendation facilities for multi-dimensional and interactive browsing of items. Exploiting user preferences, the *Prefiltering* module selects a set of *candidate* items for recommendation; successively, the *Objects Ranking* module assigns a ranking of such candidates exploiting some *ranking functions* defined on the MSN. Finally, the *Postfiltering* module dynamically selects on the base of some constraints (e.g. the item that a user is currently watching and context information) a subset of candidates.

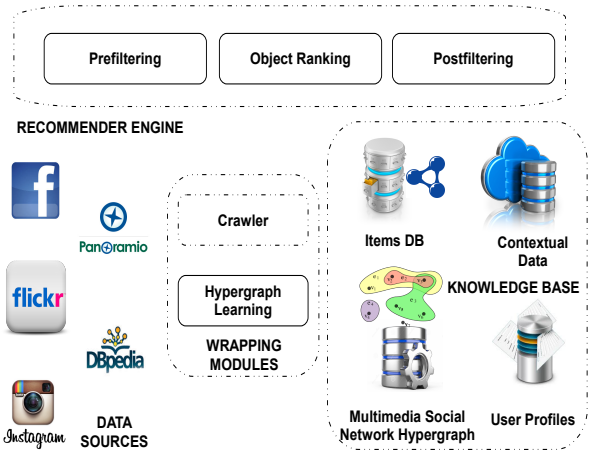


Figure 1. System Overview.

2.2 Recommendation Process

2.2.1 Pre-filtering Stage using user preferences

In the *pre-filtering* stage, our aim is to select for a given user u_h a subset $O_h^c \subset O$ containing items that are good "candidates" to be recommended.

Each item subjected to recommendation may be represented in different and heterogeneous feature spaces. For instance, a picture may be described by a set of metadata as title, description, tags, by the position in which was token and so on. Each of these sets of features contributes to the characterization of the items to different *extents*.

The first step consists in clustering together "similar" items, where the similarity should consider all (or subsets of) the different spaces of features. To this purpose, we employ *high-order star-structured co-clustering* techniques - that some of the authors have adopted in previous work [14, 7, 8] - to address the problem of heterogeneous data pre-filtering.

Let $O = \{o_1, \ldots, o_n\}$ be the set of items and $\mathcal{F} = \{F^1, \ldots, F^l\}$ a set of l feature spaces. In our recommendation problem, a user u_h is represented as a set of vectors in the same l feature spaces describing the items. To provide a first candidate list of items to be recommended, we measure the *cosine distance* of the user vectors associated to the k-th space, with the centroids of each item clusters in the k-th space. For each space, the most similar item cluster is chosen leading to l clusters $\{X_1^c, \ldots, X_l^c\}$ of candidate items.

Then, two different strategies can be adopted to provide the pre-filtered list of candidate items O_h^c: (i) *set-union strategy* - the items belonging to the union of all clusters are retained, i.e., $O_h^c = \bigcup_k X_k^c$; (ii) *threshold strategy* - the items that appears in at least ths clusters ($ths \in \{1 \ldots l\}$) are retained.

2.2.2 Ranking Stage via hypergraph modeling

The main goal of this stage is to automatically rank the set of items O embedding in a collaborative learning context: the MSN deriving by the integration of the different multimedia data sources. In particular, we use a novel technique that the authors have proposed in a previous work [6]. In our vision, a MSN is basically composed by three different kinds of entities (nodes):

- *Users* - the set of persons and organizations constituting the particular social community: several information concerning their profile, interests, preferences, etc. can eventually be considered and exploited by our model;
- *Multimedia Objects* - the set of multimedia resources (i.e. images, video, audio, posts, documents, etc.) that can be shared within a MSN community: high level (*metadata*) and low level information (*features*) can be properly used in our model;
- *Annotation Assets* - each set of symbols (e.g., keyword, tag, label, etc.) exploited by users to annotate multimedia resources within a MSN; we explicitly note that it is possible to relate a given as-

3

set with a specific *concept* (as an example a topic, a named entity, etc. which definition can be found into dictionaries, ontologies and so on), thus formally providing the related semantics.

Several types of relationships can be established among the described entities: a user can annotate an object with a particular tag, two friends can comment the same object, a user can tag another user in a photo, a user can share an object within a group, etc. In particular, we distinguish between *user to user relationships*, describing user actions towards other users, and *user to multimedia relationships*, describing user actions on objects, eventually involving some annotation assets. In addition, *similarity relationships* can be added between two objects (using multimedia features) or between two assets (by taxonomic distances).

Due to the variety and complexity of these relationships, we leverage the *hypergraph* formalism to model a MSN (all the details are provided in [6]). Then, we introduce some functions can be profitably used to "rank" users or multimedia objects in a MSN.

In our model the concept of *rank* of a given node is related to the concept of *influence*, and in our vision it can be measured by the number of user nodes that are "reachable" within a certain number of steps using any hyperpath, with respect to a social community of users, and eventually to a given *topic* of interest.

The final goal is to compute the ranking of the multimedia items in O_h^c, using as measure the social influence of each object withih the users' community.

2.2.3 Post-Filtering Stage by context information

In this stage, we have introduced a *post-filtering* method for generating the final set of "real" candidates for recommendation using *context* information.

The context is represented by means of the well-known *key-value* model [1] using as dimensions some of the different feature spaces related to items. In our system, context features can be expressed either directly using some *target items* (e.g. objects that have positively captured user attention) or specifying the related values in the shape of *constraints* that recommended items have to satisfy.

Assume that a user u_h is currently interested in a target item o_j. We can define the set of candidate recommendations as follows:

$$O_{h,j}^f = \bigcup_{k=1}^{M} \{o_i \in O_h^c \mid a_{ij}^k > 0\} \cup \{o_i \in NNQ(o_j, O_h^c)\} \tag{1}$$

The set of candidates includes the items that have been accessed by at least one user within k steps from o_j, with k between 1 and M, and the items that are most similar to o_j according to the results of a *Nearest Neighbor Query* ($NNQ(o_j, O_h^c)$) functionality. Note that a positive element a_{ij}^k of A^k indicates that o_i was accessed exactly k steps after o_j at least once. The ranked list of recommendations is then generated by ranking the items in $O_{h,j}^f$, for each item o_j selected as interesting by user u_h, using the ranking vector R_h thus obtaining the final set O_h^f.

Finally, for each user all the items that do not respect possible context connstraints are removed from the final list.

3 A Case Study

We have opportunely customized our system in order to provide touristic multimedia guiding services for users that are interested in visiting the old town center of Naples, Italy. On the base of user preferences and actual position, a set of POIs are shown on a proper map to tourists correlated with a multimedia description.

For instance, when a user is approaching a particular cultural POI (e.g. *Piazza del Gesu' Nuovo*), the related multimedia description and the set of candidate objects (i.e. images and texts related to the near POIs) are delivered on the user's mobile device (pre-filtering stage).

The list of proposed objects depends on the user's preferences (e.g. the majority of items will be images if a user prefers to see such kinds of data and will reveal effective user needs), is initially ordered according to effective user location (i.e. the closest items will appear at the top of list) and contains data grouped by the related cultural POI. Successively, after the user has selected one or more objects (for example the item he is currently watching), the recommendation services first perform a ranking (ranking stage) of all the candidate objects according to their recommendation grades and then filters the recommendation list considering only the most similar items to target objects (post-filtering stage).

When a user is near to a different POI, he/she can decide to modify the list of target objects (e.g. removing those related to the previous visited POI or adding new objects) and consequently recommendations will be automatically updated, thus including new items.

The design choices are briefly reported in the following.

- We consider as data source *Flickr*, *Instagram*, *Panoramio*, *DBPedia* and other domain digital

libraries, collecting about 500,000 items (images and texts related to historical buildings, churches, famous square and other attractions) and about 5,000 user profiles.

- As items' metadata, we consider for each multimedia item information related to *title*, *description*, *type*, *kind*, *language*, *tags*, *keyword*, *comments*, *ratings* (and for pictures the geographic position in which were taken). In addition, images are also described by a set of low-level features (i.e. SURF).

- For each item, available users' preferences, comments, feedbacks and other actions have been captured, also exploiting correlated public information from Social Networks (i.e. *Facebook*).

For what implementation details concern, the Wrapping modules leverage proper API and JAVA libraries to collect the different information of interest.

The Knowledge Base, realized using different technologies, allows to manage all the different kind of information: Contextual Data instances (messages containing information about users' position) are managed by the *Cassandra* DBMS, Items' descriptions are stored in the *Turtle* format and managed by the *AllegroGraph* repository (semantics of data can be specified by linking values of some attributes to some available ontological schema), User Profiles and the MSN hypergraph are respectively managed by *MongoDB* and *Neo4j* DBMSs.

On the other hand, the Recommender Engine exploits proper JAVA libraries (some developed for the system presented in [5] and integrated with co-clustering libraries [7] and the rank refining procedure[1]) to accomplish its tasks.

Finally, a user can interact with our system using at the moment an Android Multimedia Guide App exploiting *Google Map* API.

4 Experimental Results

Recommender Systems are very complex applications that are based on a combination of several models, algorithms and heuristics. This complexity makes evaluation efforts very difficult and thus results are hardly generalizable, as reported in the literature [3]. Moreover, characterizing and evaluating the quality of a user's experience and subjective attitude toward the acceptance of recommender technology is an important issue which we will consider in the following.

The majority of research efforts on recommender system evaluation have mainly focused on prediction *accuracy* and *stability* (e.g., [3]).

More recently, researchers began examining issues related to users subjective opinions and developing additional criteria to evaluate recommender systems. In particular, they suggest that user satisfaction does not always (or, at least, not only) correlate with the overall recommenders accuracy.

Starting from these considerations and based on current trends in the literature, we decided to perform a *user-centric* evaluation based on *user satisfaction* with respect to assigned activities, evaluating how our recommendations can effectively support browsing tasks of different complexity when the complexity of desired itesm increases.

As in our previous work [4, 5, 7, 8], we evaluate the impact of the proposed system on users engaged in several *search tasks* of multimedia items and compared its performances with the well-known *Panoramio* system[2] that, in turn, provides basic search mechanisms.

In particular, our goal was to establish how helpful our system is in assisting the search of specific multimedia objects (images) and guiding the users towards information which satisfy their interests. The dataset used in these experiments is a subset of about 10,000 items related to specific POIs.

In order to evaluate the impact of the system on the users, we have conducted the following experiments. The system was make available to a set of 50 users. These users were all interested in the cultural heritage domain, they already had experience in the use of PCs and electronic devices, even if they were not experts in ICT. We asked these users to browse the collection of items and complete several search tasks (20 tasks per user) of different complexity (five tasks for each complexity level), using Panoramio facilities. After this test, we asked them to browse the same collection with the assistance of our recommender system and complete other 20 tasks of similar complexity. We have subdivided browsing tasks in the following four broad categories:

1. ***Low Complexity*** search tasks (T_1): e.g. find at least 30 images related to 3 different POIs depicting ancient churches;

2. ***Medium Complexity*** search tasks (T_2): e.g. find at least 50 images related to 5 different POIs depicting ancient churches, historical building and famous squares (10 objects for each subject);

[1] we use *LIRE* for the content-based image retrieval

[2] http://www.panoramio.com/

5

Table 1. Comparison between our system and Panoramio in terms of t_a and n_c average values

Task Class	System	$t_a(sec)$	n_c
Low Complexity	Recommender	155	38
Low Complexity	Panoramio	164	42
Medium Compl.	Recommender	335	84
Medium Compl.	Panoramio	402	104
High Complexity	Recommender	1156	298
High Complexity	Panoramio	1302	334
Very High Compl.	Recommender	1645	351
Very High Compl.	Panoramio	1832	410

3. ***High Complexity*** search tasks (T_2): e.g. find at least 100 images related to 10 different POIs (near to the actual user position) depicting ancient churches, historical building and famous squares (10 objects for each subject);

4. ***Very High Complexity*** search tasks (T_2): e.g. find at least 150 images objects related to 10 different POIs (near to the actual user position) depicting ancient churches, historical building and famous squares (15 objects for each subject).

Note that the complexity of a task depends on several factors: the number of items to explore, the type of desired features and the number of additional constraints. Two strategies were used to evaluate the results of this experiment: (i) empirical measurements of access complexity in terms of *mouse clicks* and *time*; (ii) TLX (*NASA Task Load Index factor*).

With respect to the first strategy, we measured the following parameters: (i) *access time (t_a)* – the average time spent by the users to request and access all the images for a given class of tasks; (ii) *number of clicks (n_c)* – the average number of clicks necessary to collect all the requested images for a given class of tasks.

Table 1 reports the average values of t_a and n_c for both Panoramio and our system (Recommender), for each of the four task complexity levels defined. Especially for the most complex tasks, our system shows better performances than Panormaio, especially for the more complex tasks.

We then asked the same group of users to express their opinion about the capability of Panoramio and our system respectively to provide an effective user experience in completing the assigned search tasks, based on the TLX evaluation protocol [12].

Specifically, TLX is a multi-dimensional rating procedure that provides an overall workload score based on a weighted average of ratings on six sub-scales: mental demand, physical demand, temporal demand, own performance, effort and frustration. Lower TLX scores are better and the average scores are then reported in Table 2.

Table 2. Comparison between our system and Panoramio in terms of TLX factors for each category of users

TLX factor	Recommender	Panoramio
Mental demand	39	41
Physical demand	36.3	48
Temporal demand	39	50
Effort	35	50.5
Perfomances	69.7	79.8
Frustation	33.2	44.1

Our system outperforms in a significative way Panormaio in every sub-scale except for *mental demand* and *performance*: this happens because sometimes an expert user considers the automatic suggestions not useful, just because they know what they are looking for.

In summary, our system provides a better (less frustrating) user experience during the search tasks. In addition, the fact that search tasks can be completed faster using our system is an indication that recommendations are effective, as they allow a user to explore interesting and related items one after another, without the interference of undesired items that would otherwise slow down the process.

5 Conclusions and Future Work

In this paper a novel multimedia and scoial recommendation approach for Cultural Heritage applications. It combines several aspects of users - i.e. their *preferences* (usually in the shape of items' metadata) and *interactions* within a social community modeled using hypergraphs - together with items' *multimedia features* and *context information* within a general framework that can support different applications (touristic guiding services for museums, visiting paths recommendation for old town centers and archeological sites, etc.).

Preliminary experiments on user satisfaction demonstrated how our approach achieve very promising and interesting results. Future works will be devoted to extend the experimental evaluation to a larger multimedia data set, also considering the performance, evaluated in terms of accuracy, precision

6

and recall, of the performed recommendations. Moreover, we plan to apply our approach to other kinds of data gathered from heterogeneous collections and compare our approach with other ones proposed in the literature.

References

[1] Gediminas Adomavicius, Ramesh Sankaranarayanan, Shahana Sen, and Alexander Tuzhilin. Incorporating contextual information in recommender systems using a multidimensional approach. *ACM Transactions on Information Systems (TOIS)*, 23(1):103–145, 2005.

[2] Gediminas Adomavicius and Alexander Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17:734–749, 2005.

[3] Gediminas Adomavicius and Jingjing Zhang. Stability of recommendation algorithms. *ACM Transactions on Information Systems (TOIS)*, 30(4):23, 2012.

[4] Massimiliano Albanese, Antonio d'Acierno, Vincenzo Moscato, Fabio Persia, and Antonio Picariello. A multimedia semantic recommender system for cultural heritage applications. In *Semantic Computing (ICSC), 2011 Fifth IEEE International Conference on*, pages 403–410. IEEE, 2011.

[5] Massimiliano Albanese, Antonio dAcierno, Vincenzo Moscato, Fabio Persia, and Antonio Picariello. A multimedia recommender system. *ACM Transactions on Internet Technology (TOIT)*, 13(1):3, 2013.

[6] Flora Amato, Vincenzo Moscato, Antonio Picariello, and Giancarlo Sperli'. Multimedia social network modeling: a proposal. In *Semantic Computing (ICSC), 2016 10th IEEE International Conference on*. IEEE, 2016.

[7] Ilaria Bartolini, Vincenzo Moscato, Ruggero G Pensa, Antonio Penta, Antonio Picariello, Carlo Sansone, and Maria Luisa Sapino. Recommending multimedia visiting paths in cultural heritage applications. *Multimedia Tools and Applications*, pages 1–30, 2014.

[8] Francesco Colace, Massimo De Santo, Luca Greco, Vincenzo Moscato, and Antonio Picariello. A collaborative user-centered framework for recommending items in online social networks. *Computers in Human Behavior*, 51:694–704, 2015.

[9] Ruihai Dong, Michael P. O'Mahony, Markus Schaal, Kevin McCarthy, and Barry Smyth. Sentimental product recommendation. In *Proceedings of the 7th ACM Conference on Recommender Systems*, RecSys '13, pages 411–414, New York, NY, USA, 2013. ACM.

[10] Paul Dourish. What we talk about when we talk about context. *Personal and ubiquitous computing*, 8(1):19–30, 2004.

[11] Gayatree Ganu, Yogesh Kakodkar, and AméLie Marian. Improving the quality of predictions using textual information in online user reviews. *Inf. Syst.*, 38(1):1–15, March 2013.

[12] Sandra G Hart and Lowell E Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Advances in psychology*, 52:139–183, 1988.

[13] Yoshinori Hijikata, Kazuhiro Iwahama, and Shogo Nishida. Content-based music filtering system with editable user profile. In *Proceedings of the 2006 ACM symposium on Applied computing*, SAC '06, pages 1050–1057, New York, NY, USA, 2006. ACM.

[14] Dino Ienco, Céline Robardet, Ruggero G Pensa, and Rosa Meo. Parameter-less co-clustering for star-structured heterogeneous data. *Data Mining and Knowledge Discovery*, 26(2):217–254, 2013.

[15] Katerina Kabassi. Personalisation systems for cultural tourism. In *Multimedia services in intelligent environments*, pages 101–111. Springer, 2013.

[16] Alexandros Karatzoglou, Xavier Amatriain, Linas Baltrunas, and Nuria Oliver. Multiverse recommendation: n-dimensional tensor factorization for context-aware collaborative filtering. In *Proceedings of the fourth ACM conference on Recommender systems*, pages 79–86. ACM, 2010.

[17] Yehuda Koren. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '08, pages 426–434, New York, NY, USA, 2008. ACM.

[18] Cane WK Leung, Stephen CF Chan, and Fu-lai Chung. Integrating collaborative filtering and sentiment analysis: A rating inference approach. In *Proceedings of The ECAI 2006 Workshop on Recommender Systems*, pages 62–66. Citeseer, 2006.

[19] Veronica Maidel, Peretz Shoval, Bracha Shapira, and Meirav Taieb-Maimon. Evaluation of an ontology-content based filtering method for a personalized newspaper. In *Proceedings of the 2008 ACM conference on Recommender systems*, RecSys '08, pages 91–98, New York, NY, USA, 2008. ACM.

[20] Katarzyna Musial, Krzysztof Juszczyszyn, and Przemyslaw Kazienko. Ontology-based recommendation in multimedia sharing systems. *System Science*, 34:97–106, 2008.

[21] Nikolaos Pappas and Andrei Popescu-Belis. Sentiment analysis of user comments for one-class collaborative filtering over ted talks. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '13, pages 773–776, New York, NY, USA, 2013. ACM.

[22] Michael Pazzani and Daniel Billsus. Content-Based Recommendation Systems. pages 325–341. 2007.

[23] Naren Ramakrishnan, Benjamin J. Keller, Batul J. Mirza, Ananth Y. Grama, and George Karypis. Privacy risks in recommender systems. *IEEE Internet Computing*, 5:54–62, November 2001.

[24] Paul Resnick, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom, and John Riedl. Grouplens: An open architecture for collaborative filtering of netnews. pages 175–186. ACM Press, 1994.

[25] Francesco Ricci, Lior Rokach, Bracha Shapira, and Paul B. Kantor, editors. *Recommender Systems Handbook*. Springer, 2011.

[26] Andrew I. Schein, Alexandrin Popescul, Lyle H. Ungar, and David M. Pennock. Methods and metrics for cold-start recommendations. In *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '02, pages 253–260, New York, NY, USA, 2002. ACM.

[27] Vivek Kumar Singh, Mousumi Mukherjee, and Ghanshyam Kumar Mehta. Combining collaborative filtering and sentiment classification for improved movie recommendations. In *International Workshop on Multi-disciplinary Trends in Artificial Intelligence*, pages 38–50. Springer, 2011.

[28] Johann Stan, Fabrice Muhlenbach, Christine Largeron, et al. Recommender systems using social network analysis: Challenges and future trends. *Encyclopedia of Social Network Analysis and Mining*, pages 1–22, 2014.

[29] Xiaoyuan Su and Taghi Khoshgoftaar. A survey of collaborative filtering techniques. *Advances in Artificial Intelligence*, 2009, 2009.

[30] Hilmi Yildirim and Mukkai S. Krishnamoorthy. A random walk method for alleviating the sparsity problem in collaborative filtering. In *Proceedings of the 2008 ACM conference on Recommender systems*, RecSys '08, pages 131–138, New York, NY, USA, 2008. ACM.

[31] Hsiang-Fu Yu, Cho-Jui Hsieh, Si Si, and Inderjit S Dhillon. Parallel matrix factorization for recommender systems. *Knowledge and Information Systems*, pages 1–27, 2013.

[32] Xujuan Zhou, Yue Xu, Yuefeng Li, Audun Josang, and Clive Cox. The state-of-the-art in personalized recommender systems for social networking. *Artif. Intell. Rev.*, 37(2):119–132, February 2012.

8

VisCFSM: Visual, Constraint-Based, Frequent Subgraph Mining

Jennifer L. Leopold
Missouri University of Science & Technology
Department of Computer Science
Rolla, Mo USA
leopoldj@mst.edu

Chaman L. Sabharwal
Missouri University of Science & Technology
Department of Computer Science
Rolla, MO USA
chaman@mst.edu

Nathan W. Eloe
Northwest Missouri State University
School of Computer Science and Information Systems
Maryville, MO USA
nathane@nwmissouri.edu

***Abstract*—Graphs long have been valued as a pictorial way of representing relationships between entities. Contemporary applications use graphs to model social networks, protein interactions, chemical structures, and a variety of other systems. In many cases, it is useful to detect patterns within graphs. For example, one could be interested in identifying frequently occurring subgraphs, which is known as the frequent subgraph mining problem. A complete solution to this problem can result in numerous subgraphs and can be time-consuming to compute. An approximate solution is faster, but is subject to static heuristics that are beyond the control of the user. Herein we present VisCFSM, a visual, constraint-based, frequent subgraph mining system which allows the user to dynamically specify a variety of constraints on the subgraphs to be found while the mining algorithm is running. The constraint specification interactions are performed through a visual user interface, thereby facilitating a form of visual algorithm steering. This approach can be integrated with any frequent subgraph mining algorithm. Most importantly, this approach has the potential for the user to better, and more quickly, find the information that is of most interest to him/her in a graph.**

Keywords-graph; data mining; visual algorithm steering

I. INTRODUCTION

Graphs long have been valued as a pictorial way of representing complex relationships between entities. Commercial, research, and government organizations use graphs to model social networks, protein interactions, chemical structures, and a variety of other systems. A common application of graph data mining is to identify the most recurrent relationships or patterns amongst the data in a graph, which typically requires finding frequently occurring subgraphs.

For some applications, the input will be a collection of relatively small graphs, and the search for frequent subgraphs is performed over each individual graph in the collection before those results are combined. This is known as a *graph-transaction setting*. In contrast, the input may be a single graph; this is referred to as a *single graph setting*. Our work refers to the latter environment. We also restrict our work to *static graphs*, and do not address *dynamic graphs* or *streaming graphs*, which are discussed in [1].

DOI reference number: 10.18293/DMS2016-026

Formally, we define the Frequent Subgraph Mining (FSM) problem as in the paper by Abedijaberi [2] using Definitions 1-4 given below.

Definition 1. *A labelled graph $G = (V, E, L_V, L_E)$ consists of a set of vertices V, a set of undirected or directed edges E, and two labeling functions L_V and L_E that association labels with vertices and edges, respectively.*

It should be noted that the labels of any two vertices (or any two edges) may not be unique. However, each vertex (and each edge) will have a unique *id*.

Definition 2. *A graph $S = (V_S, E_S, L_{VS}, L_{ES})$ is a subgraph of $G = (V, E, L_V, L_E)$ iff $V_S \subseteq V$, $E_S \subseteq E$, $L_{VS}(v) = L_V(v)$ and $L_{ES}(e) = L_E(e)$ for all $v \in V_S$ and $e \in E_S$.*

Definition 3. *A subgraph isomorphism of S to G is a one-to-one function $f: V_S \rightarrow V$ where $L_{VS}(v) = L_V(f(v))$ for all vertices in $v \in V_S$, and for all edges $(u,v) \in E_S$, $((f(u),f(v)) \in E$ and $L_{ES}(u,v) = L_E((f(u), f(v))$.*

Definition 4. *Let I_S be the set of isomorphisms of a subgraph S in graph G. Given a minimum support threshold τ, the frequent subgraph mining problem (FSM) is to find all subgraphs S in G such that $|I_S| \geq \tau$.*

The advantage of limiting frequent subgraphs to only those with disjoint edges is computational tractability [3]. But this comes at the expense of disregarding potentially useful information. Hence, in our work we allow isomorphic subgraphs to share edges.

FSM algorithms that find complete solutions may, depending upon the specified threshold value and the size of the graph, result in numerous subgraphs and take a considerable amount of time to compute. Algorithms that find approximate solutions are faster, but apply static heuristics that are beyond the control of the user (unless s/he modifies the software).

Herein we present VisCFSM, a visual, constraint-based, frequent subgraph mining system which allows the user to dynamically specify a variety of constraints on the subgraph mining algorithm while it is running. The constraint specification interactions are performed through a visual user interface, thereby facilitating a form of visual algorithm steering.

The prototype implementation uses the FSG [4] frequent subgraph algorithm; however, the approach we employ can be integrated with any FSM algorithm. Most importantly, this approach has the potential for the user to better, and more quickly, find the information that is of most interest to him/her in a graph.

The organization of this paper is as follows. Section II provides a brief overview of related work in graph data mining. Motivation for the need for dynamic, visual steering of FSM using constraints is presented in Section III. In Section IV, we discuss the VisCFSM infrastructure in terms of the FSM, the constraint satisfaction system, and the graphical user interface. An example of running VisCFSM is presented in Section V. Finally, we discuss our plans for future work in Section VI and conclusions in Section VII.

II. RELATED WORK

A. Graph Data Mining Algorithms

Graph Data Mining (GDM) algorithms are divided into three main categories: Graph Theory Based, Inductive Logic Programming, and Greedy Search [5]. Our work focuses on the Graph Theory Based category, which consists of two main groups: Apriori-based and pattern growth-based approaches. Algorithms in the first group generate candidate subgraphs by joining two frequent subgraphs of the same size to generate larger subgraphs. Pattern growth algorithms generate candidates by adding a new edge to each smaller frequent subgraph.

FSM algorithms typically face two computational challenges: (i) candidate subgraph generation, and (ii) identification of candidate subgraphs that meet the minimum support threshold. In the worst case, all subgraphs in the graph must be examined, which is exponential in complexity, and subgraph isomorphisms must be computed, which is an NP-complete problem. FSM algorithms may attempt to improve runtime performance by reducing the size of the search space, avoiding duplicate comparisons, and/or minimizing the amount of memory required for compiling intermediate results. Another solution to reduce the runtime is to provide an approximate, rather than a complete, solution to the FSM problem.

B. Heuristics for Approximate Solutions

Heuristic FSM algorithms such as SUBDUE [6], GREW [7] and GRAMI [8] discover only a subset of all frequent subgraphs of a graph. These algorithms do not return any infrequent patterns (i.e., the results do not have false positives), but may miss some frequent ones (i.e., the results effectively may have false negatives). The type of heuristics that are employed are quite diverse, and also vary considerably in their degree of complexity. Some examples are listed below:

- SUBDUE [6] starts with frequent subgraphs consisting of a single vertex, and then expands those in a breadth-first manner by adding a new edge. The order of processing is known as a “beam search”, and only a predetermined number of paths (i.e., the beam width) are kept as candidates at each iteration. Hence some valid frequent subgraphs will be missed.
- Like SUBDUE, GREW [7] employs a beam search to prune large portions of the search space. It also iteratively joins frequently occurring pairs of nodes into a single supernode, and determines disjoint embeddings of connected subgraphs using a maximal independent set algorithm. GREW employs an additional heuristic that deliberately underestimates the frequency of each discovered subgraph in an attempt to reduce the search space. While experiments showed that GREW significantly outperformed SUBDUE with respect to runtime, those experiments showed that this came at the expense of finding fewer frequent subgraphs.
- Pattern growth algorithms generate candidate subgraphs by adding a new edge to smaller frequent subgraphs. GRAMI [8] only adds frequently occurring edges to smaller frequent subgraphs when generating candidate subgraphs. This will miss finding some valid frequent subgraphs, but reduces the total number of iterations over edges that must be considered.
- AGRAMI [8] is an extension of GRAMI that employs additional heuristics in an effort to scale to larger graphs. For example, it enforces a timeout when testing whether a subgraph occurs at least as many times as the minimum support threshold; if the solution cannot be computed within a particular amount of time, that subgraph is assumed to be infrequent.

In the same paper that presents GRAMI and AGRAMI [8], the authors briefly discuss CGRAMI, a version of GRAMI that seeks to find more general patterns in graphs than just frequent subgraphs. This work is noteworthy to mention herein because it claims to support the following user-defined constraints:

- Number of vertices (or edges) in a pattern cannot exceed a specified value
- Vertex degree in a pattern cannot exceed a certain value
- A pattern must include/exclude only vertices with certain labels
- A pattern must include only certain edges
- A pattern cannot include certain edges
- A pattern cannot include a specified subgraph
- A specified vertex label cannot appear more than N times in a pattern

To specify desired constraints in CGRAMI, the user must comment out certain lines of code (and uncomment other lines) for the constraints, set the values for parameters, and then recompile the program. The program has a command-line interface; there is no graphical user interface.

As stated previously, what all heuristic FSM (and constraint-based GDM) algorithms have in common is the inability for the user to dynamically customize the heuristics, or any form of constraints, while the algorithm is running. This is the novel contribution of the work presented herein.

III. MOTIVATION

In part, motivation for this work came from a graduate course on Advanced Data Mining taught by author Leopold in 2015 at Missouri University of Science and Technology. That year the focus of the course was graph data mining. Students read several research papers on GDM algorithms and applications. Some of the students in the class implemented a few of the algorithms in Python, but were frustrated that they had to wait a considerable amount of time for the computation on some relatively small graphs (e.g., a graph of 50 vertices with average vertex degree 3.5 took over 8 hours to compute all frequent subgraphs of minimum support 2). When they had their programs output intermediate results as the subgraphs were being found, sometimes the students would terminate the program, and restart it with a different threshold value to further discriminate the result set and make the program finish more quickly.

At the end of the course, the students were asked what kinds of constraints they would have found useful to "steer" a FSM algorithm dynamically, even if it meant that the resulting set of subgraphs would not be complete. Here we use the term *steering* as discussed in [9]: the ability to have a continuous visualization of the (output) data as a program executes, coupled with the ability for the programmer to interactively modify any aspect of the program and see the effects without restarting the computation.

With a social network (specifically, a terrorist network) as an application domain, the students identified the constraints and use case examples listed below. In this social network, it is assumed that a vertex in the graph is labelled with a person's name, which are not necessarily unique. Additional information about a person and his/her relationship to other people may be represented in the graph as vertex or edge data.

- **Include/exclude subgraphs containing a certain set of vertices.** Ex.: Suppose that we're using a social media network to identify terrorist threats. The number of frequent subgraph results may be quite high at first due to very small terrorist groups. So we then want to narrow our search, and only continue to look for subgraphs that include a specific group of people that we know conduct terrorist activities.
- **Include/exclude only frequent subgraphs that appear more/less than subgraph (or vertex) X does.** Ex.: We see a specific name in the preliminary results of our search that we already know is a leader and a threat. But his name isn't the only one we see, and we want to know who in the group is more important, of high rank, or higher rank than this person. So we then narrow our search to find subgraphs that appear more often than those containing this person. Or maybe we are looking for someone we can capture and get information from, in which case we look for someone important who appears less often.
- **Include/exclude only frequent subgraphs that are disconnected/connected to subgraph (or vertex) X.** Ex.: We begin a search on terrorist cells. However, based on seeing a particular group appearing frequently in the results, we want to narrow our search to those connected to that group. Similarly, if we are trying to identify new terrorist cells or rival cells, we may want to look only at those groups that are disconnected from a certain group.
- **Include/exclude only frequent subgraphs where the average vertex degree is greater than some number.** Ex.: We're looking for potential terrorist cells, and not interested in groups with only a couple of connections; such groups are unlikely to be funded or be a real threat. We may not see this until after we have seen the initial (small-sized) frequent subgraphs.
- **Include/exclude subgraphs containing a certain number of edges.** We may not be interested in seeing small terrorist groups, but rather want to see a certain amount of interconnectivity; these might prove to be the more dangerous terrorist groups.
- **Change minimum support.** Ex.: We may start our search very wide open, but, after seeing some preliminary results that are too numerous and/or contain trivial information (e.g., everyone is a potential terrorist), decide that we want to raise the threshold.

It should be noted that these constraints are not intended to be mutually exclusive, but rather conjunctive; we should be able to specify any combination of constraints.

In the next section, we discuss the VisCFSM infrastructure in terms of the FSM, the constraint satisfaction system, and the graphical user interface. The system was designed to address many of the above listed constraints.

IV. VISCFSM

The infrastructure of VisCFSM consists of a front end and a back end. The front end is comprised of the graphical user interface which displays the frequent subgraphs as they are computed, and allows the user to visually steer the FSM by specifying constraints on frequent subgraph selection as the algorithm is progressing. The back end consists of the FSM and the constraint satisfaction system. In this section we briefly discuss each part of the infrastructure.

A. The FSM

As mentioned in Section II, we have chosen to focus on the Graph Theory Based category of graph data mining algorithms, which consists of Apriori-based and pattern growth-based approaches. For the prototype implementation of VisCFSM we chose a pattern growth algorithm, FSG [4]. The algorithm starts by finding all frequent subgraphs consisting of one edge. It then makes repeated iterations, generating candidates by adding a new edge to each of the largest frequent subgraphs found so far. This particular algorithm was selected primarily for its simplicity; it is certainly not one of the most efficient FSM algorithms that exists, but we believed that the logic upon which it is based could easily be understood by most users. The choice of FSM algorithms to be used in VisCFSM is not important; the constraint satisfaction system and visualization control system that we employ actually can be integrated with any FSM algorithm.

B. *The Constraint Satisfaction System*

Inspired by the use cases presented in Section III, several structural and semantic constraints have been implemented for VisCFSM. These are listed below:

- Include/exclude frequent subgraphs that contain certain vertices or edges
- Include/exclude frequent subgraphs that include a particular subgraph
- Include subgraphs that are connected/disconnected to a particular vertex or edge
- Include only frequent subgraphs that have at least one vertex that has degree greater than a specified number
- Include only frequent subgraphs where the average vertex degree greater than a specified number
- Include/exclude frequent subgraphs containing a certain number of edges
- Exclude frequent subgraphs where a certain vertex label appears greater than a specified number of times
- Change minimum support

The user interface allows the user to specify the constraints that should be applied to the set of frequent subgraphs found so far, and whether to continue applying these constraints in the next iteration of the algorithm in an effort to find new frequent subgraphs (e.g., in the case of the FSG algorithm, the next iteration adds an edge to each of the largest-sized frequent subgraphs found so far in order to form new candidate frequent subgraphs).

C. *The Graphical User Interface*

The VisCFSM FSM and constraint satisfaction system were implemented in SWI-Prolog. The choice of a logic programming language seemed most suitable for modeling a constraint satisfaction problem. However, SWI-Prolog has no graphical capabilities. Hence, the VisCFSM graphic user interface was developed in Python.

The graphic user interface (GUI) consists of the following controls: (i) a file chooser to allow the user to select a Prolog file that contains the specification of a graph, (ii) a text input field to specify the name of the graph (i.e., a Prolog file may contain multiple graph specifications, each defined as a relation), (iii) a text input field to specify the minimum support threshold for considering a subgraph to be frequent, (iv) a constraint editor, (v) a control button to start the FSG by finding the smallest-sized FSGs, and (vi) a control button to add an edge to each of the largest FSGs found thus far. A graph specification consists of a Prolog list containing the list of vertices (in the format [ID, label]) and a list of edges, where each edge is represented as a list of two vertices. Fig. 1 shows the GUI after an undirected graph named *sampleGraph* has been loaded from a Prolog file named *graph.pl*. In this figure, no frequent subgraphs have been found yet.

The constraint editor allows the user to set up rules to filter the frequent subgraphs that will be reported. Examples of the constraint editor are shown in Fig. 2 and Fig. 3. Constraints are represented in Disjunctive Normal Form (DNF); that is, as a series of AND clauses OR'd together. The editor includes a drop-down menu of the possible constraints, a text input field for specifying the arguments to a constraint, and a display of the DNF clauses that have been specified so far. Help text is also provided to guide the user in specifying the arguments correctly.

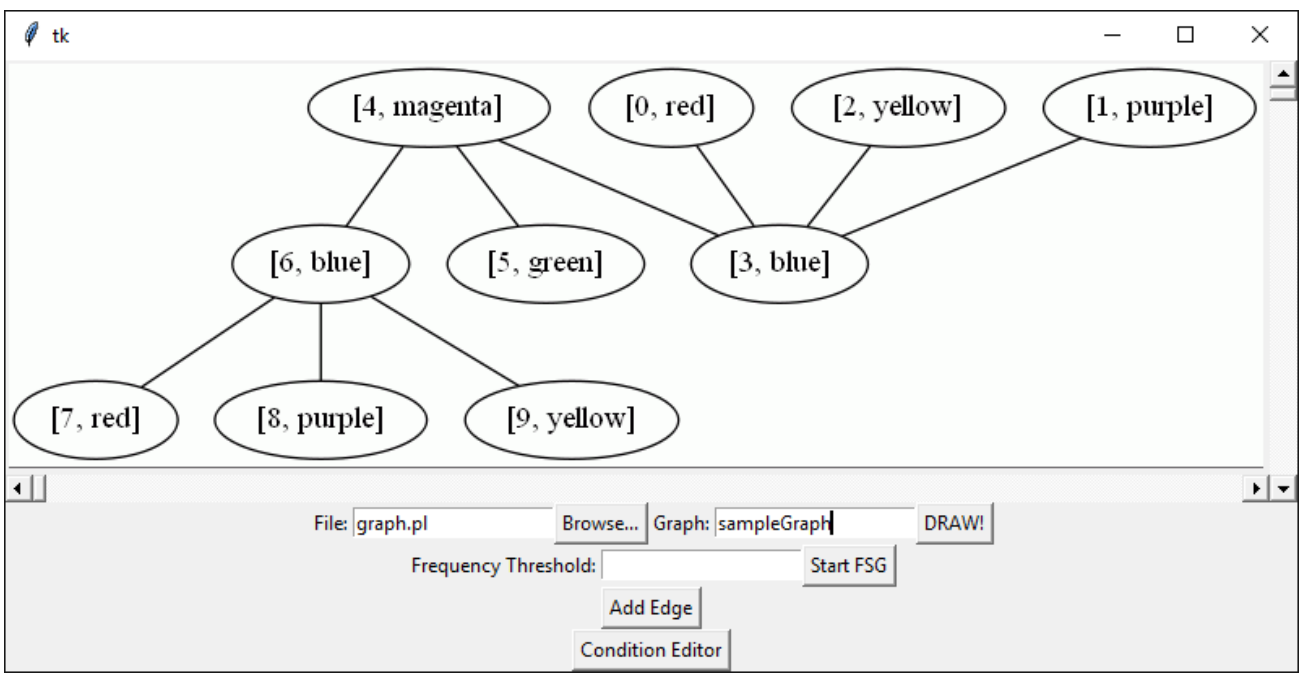


Figure 1. VisCFSM GUI after a sample graph has been loaded

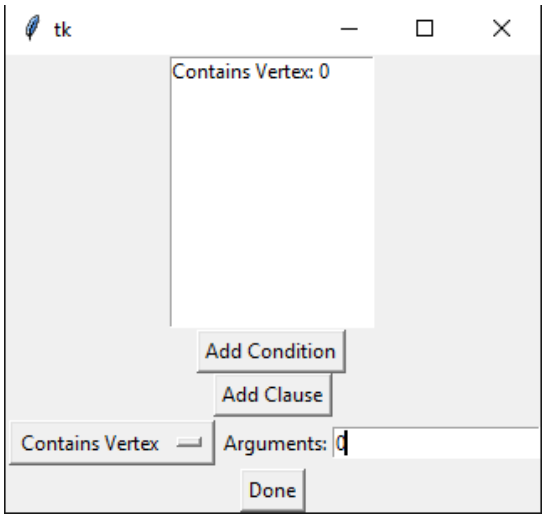


Figure 2. Constraint editor showing one clause

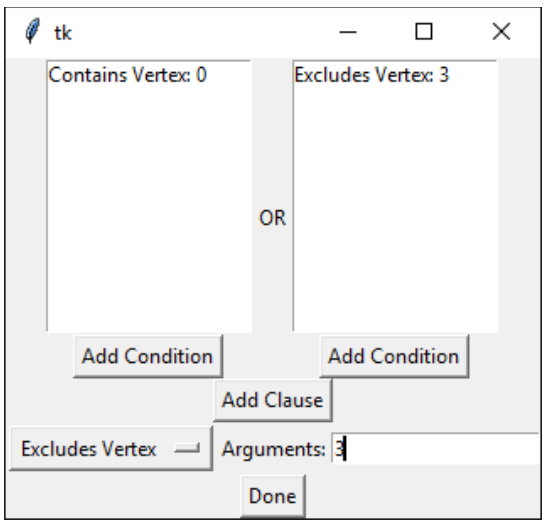


Figure 3. Constraint editor showing two clauses

The main display area in the GUI initially shows the graph that the user has specified from the selected Prolog file. Once the FSG algorithm is invoked, that area of the GUI is used to display the frequent subgraphs found in the most recent iteration of the algorithm. Recall that FSG starts by finding all frequent single-edge subgraphs, then makes repeated iterations, adding a new edge to each of the largest frequent subgraphs found so far. In future refinements of the GUI, the user will be given the ability to scroll back to previously displayed sets of (smaller-sized) frequent subgraphs, and also will be given the option to undo/redo the application of constraints and edge additions. With the current implementation, at any time, the user may restart the FSG generation algorithm from the beginning by clicking on the **Start FSG** button.

V. An Example in VisCFSM

To demonstrate the concepts behind VisCFSM, here we walk through two simple examples. We start by assuming that we have reached the state shown in Fig. 1, having specifying the file *graph.pl* and selecting *sampleGraph* as the desired graph. Clicking on the **DRAW**! button renders the graph without finding any frequent subgraphs.

For this particular graph, a minimum support threshold of 2 provides interesting results. Fig. 4 shows all frequent subgraphs with a minimum support of 2. These are listed individually with their unique vertex IDs rather than, for example, simply reporting that a subgraph with edge *(blue, red)* occurs at least 2 times; there is, however, an option in the GUI to report the results only by unique label combinations in the subgraphs.

Suppose that we want only subgraphs that contain a specific vertex, say those with the ID 0. The user would first click on the **Condition Editor** button, which will open the condition editor window. In the condition editor, the user would then select "Contains Vertex" from the dropdown menu, and specify the ID of the vertex to include; see Fig. 2. Finally, the user would click on the **Add Condition** button. The constraint editor can be closed by clicking on **Done** or closing the window.

Upon completing those steps, if the user now clicks the **Start FSG** button, the FSG algorithm will be restarted with the specified constraint applied; these results are shown in Fig. 5. At this point, the user may increment the size of these subgraphs (by clicking on the **Add Edge** button), or go back to the constraint editor, add additional constraints to the current AND clause, and recompute the set of frequent single-edge subgraphs.

To add an alternative set of constraints (i.e., add another clause that will be OR'd in the DNF representation), the user can open the constraint editor and click on the **Add Clause** button. A new clause column will be displayed, and the user now can add constraints to either of the clauses. Fig. 3 shows the addition

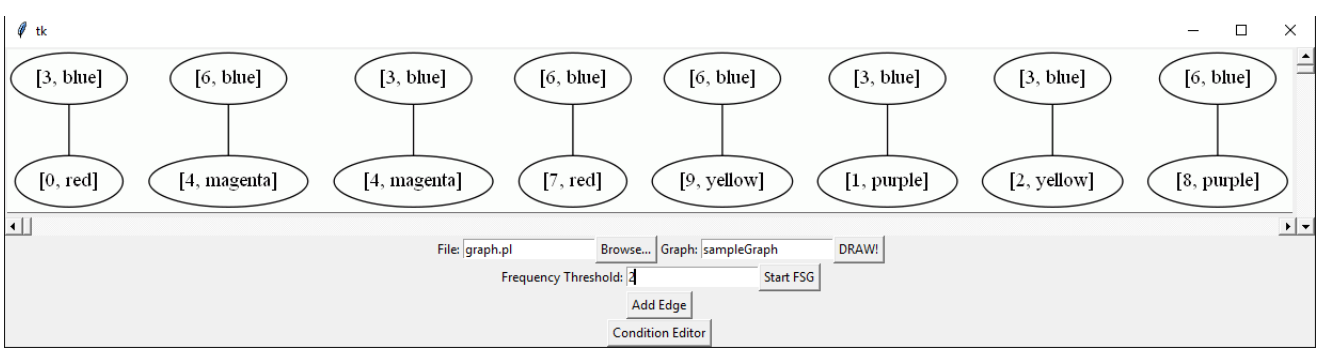


Figure 4. Subgraphs from the graph shown in Fig. 1 that occur with minimum support 2

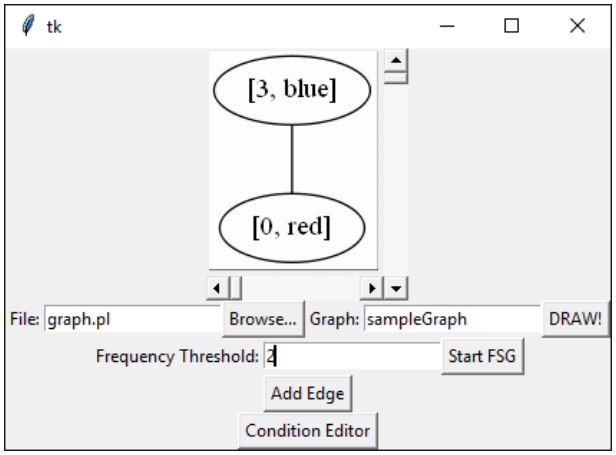


Figure 5. Results of applying the constraint specified in Fig. 2 when finding single-edge frequent subgraphs for the graph shown in Fig. 1

of another constraint in our example; namely we also want frequent subgraphs that exclude the vertex with ID 3. The resulting set of frequent single-edge subgraphs (obtained by clicking the **Start FSG** button) is shown in Fig. 6. The result set now includes single-edge subgraphs that occur at least 2 times, and include vertices with ID 0 or exclude vertices with ID 3.

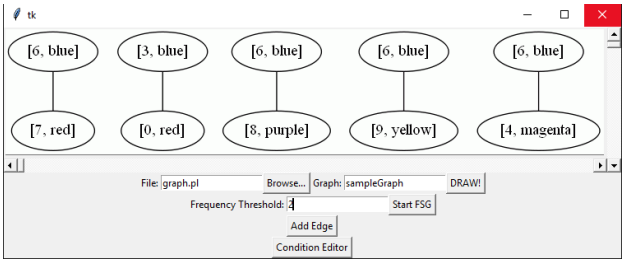


Figure 6. Results of applying the constraints specified in Fig. 3 when finding single-edge frequent subgraphs for the graph shown in Fig. 1

VI. Future Work

Future work on VisCFSM will be focused on three areas. First, we intend to perform informal studies of the usefulness and usability of the system when author Leopold again teaches the graduate course on Advanced Data Mining at Missouri University of Science and Technology in 2017. This should yield ideas for improving existing features in the user interface and adding new constraints.

Secondly, we plan to explore the incorporation of specifications of constraints by natural language and/or gestures/drawing in the user interface. The use of multimodal user interfaces for spatially-oriented applications have been of interest for years, particularly for GIS applications; see the discussion in [10], for example. Such interfaces have become a particularly active research topic in the past few years with the prevalence of mobile GIS applications (e.g., [11], [12]). Many of the same concepts for querying general spatial settings (e.g., objects such as buildings, connections between objects such as roads, directionality, etc.) should also be applicable to specifying patterns in a graph.

The third focus for future work will be improving the runtime efficiency of VisCFSM. Most algorithms used for graph data mining are designed for implementations in procedural languages. It is possible that some of our FSM operations (written in Prolog) could be sped up by expressing the steps of those algorithms (e.g., testing for graph isomorphism) in a more logical form rather than mimicking procedural solutions.

Additionally, many parts of the algorithm are embarrassingly parallel in nature; operations that are performed on every subgraph in the list of frequent subgraphs are independent from those being applied to other subgraphs, and as such can be done in a parallel manner. SWI-Prolog offers the concurrent/3 and concurrent_maplist/N predicates that automatically distribute the execution of goals across multiple threads. As threading introduces significant overhead to the process, determining the appropriate size of problems to apply concurrent operations will require empirical experimentation so as to not degrade the performance of the system.

Another option for exploring concurrency in VisCFSM includes using a Prolog interpreter embedded in the language in which the GUI is implemented. Javascript, for example, has a library that is a Prolog interpreter implemented in Javascript. This would allow the GUI to use the threading and GUI features of the host language (in Javascript's case, Electron or NW.js could provide the GUI features) while preserving the Prolog computational backend. This has the disadvantage of requiring the GUI to manage communication between threads, but the higher-level threading abilities may make this a preferable option to sending queries to a Prolog process.

If reconciling the data representations of Prolog and a host language prove to introduce more complexity than is needed, a domain-specific logic language that integrates tightly with other languages (such as miniKanren [13]) would provide a more native approach to interacting with the computational backend. This would provide access to high-level threading and GUI capabilities of more general-purpose programming languages while leveraging the constraint programming capabilities of a logic programming language.

VII. Summary and Conclusions

Herein we have presented the infrastructure for a graph mining system that provides the ability for the user to dynamically customize a variety of semantic and structural constraints while the algorithm is working to complete its overall task. Effectively, this system supports visual algorithm steering, providing the ability for the user to continuously visualize the results of the graph mining program as it executes, interactively modify aspects of the program, and see the effects without restarting the computation from the very beginning. Such capabilities are extremely valuable when dealing with graph mining, wherein the data representation is intrinsically visual, and patterns of interest may not become obvious until preliminary results are seen. Because frequent subgraph mining is a computationally intensive problem, the ability to dynamically adjust constraints on the computation can allow the user to more quickly find the information that is of most interest to him/her in a graph.

REFERENCES

[1] A. Bifet, G. Holmes, B. Bfahringer, and R. Gavald'a, "Mining frequent closed graphs on evolving data streams", Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, 2011, pp. 591-599.

[2] A. Abedijaberi and J.L. Leopold, "FSMS: a frequen subgraph mining algorithm using mapping sets," Proceedings of the 12th International Conference on Machine Learning and Data Mining, New York, NY, 2016.

[3] M. Kuramochi and G. Karypis, "Finding frequent patterns in a large sparse graph," Data Mining and Knowledge Discovery, 11.3, 2005, pp. 243-271.

[4] M. Kuramochi and G. Karypis, "Frequent subgraph discovery," Proceedings of the 2001 IEEE International Conference on Data Mining, IEEE Computer Society, 2001.

[5] M. Gholami and A. Salajegheh, "A survey on algorithms of mining frequent subgraphs," International Journal of Engineering Inventions, 1.5, 2012, pp. 60-63.

[6] D.J. Cook and L.B. Holder, "Substructure discovery using minimum description length and background knowledge, " Journal of Artificial Intelligence Research, 1(1), 1994, pp. 231-255.

[7] M. Kuramochi and G. Karypis, "An efficient algorithm for discovering frequent subgraphs," IEEE Transactions on Knowledge and Data Engineering, 16(9), September 2004, pp. 1038-1051.

[8] M. Elseidy, E. Abdelhamid, S. Skiadopoulos, and P. Kalnis, "GRAMI: frequent subgraph and pattern mining in a single large graph," Proceedings of VLDB Endowment, 2014, pp. 517-528.

[9] J.W. Atwood, Jr., M.M. Burnett, R.A. Walpole, E.M. Wilcox, and S. Yang, "Steering programs via time travel", 1996 IEEE Symposium on Visual Languages, Boulder, CO, 1996, pp. 1-8.

[10] I. Schlaisich and M.J. Egenhofer, "Multimodal spatial querying: what people sketch and talk about", 1st International Conference on Universal Access in Human-Computer Interaction, 2001, pp. 732-736.

[11] F. Cutugno, V.A. Leano, R. Rinaldi, and G. Mignini, "Multimodal framework for mobile application", Proceedings of the International Working Conference on Advanced Visual Interfaces, New York, NY, 2012, pp. 197-203.

[12] J. Doyle, M. Bertolotto, and D. Wilson, "Evaluating the benefits of multimodal interface design for CoMPASS – a mobile GIS", GeomInformatica, Volume 14, Issue 2, April 2010, pp. 135-162.

[13] C.E. Alvis, J.J. Wilcock, K.M. Carter, W.E. Byrd, and D.P. Friedman, "cKanren miniKanren with constraints", 2011.

KnotSketch: a tool for knot diagram sketching, encoding and re-generation

Gennaro Costagliola[1], Mattia De Rosa[1], Andrew Fish[2], Vittorio Fuccella[1], Rafiq Saleh[3], Sarah Swartwood[2]
[1]Dipartimento di Informatica, University of Salerno, Via Giovanni Paolo II, 84084 Fisciano (SA), Italy
[2] School of Computing, Engineering and Mathematics, University of Brighton, UK
[3]Department of Earth, Ocean and Ecological Sciences, University of Liverpool, UK
{gencos, matderosa, vfuccella}@unisa.it, {andrew.fish,S.Swartwood}@brighton.ac.uk

Abstract

Knots occur in many areas of science and art. The mathematical field of Knot Theory studies an idealised form of knots by viewing them as closed loops in 3-space. They can be formally studied via knot drawings which are well-behaved projections of the knot onto the 2-D plane. Equivalence of knots in 3-space (ambient isotopy) can be encapsulated via sequences of diagram rewriting rules, called Reidemeister moves, but finding such sequences demonstrating isotopy of two knots can be immensely challenging. Whilst there are some sophisticated tools available for some knot theoretic tasks, there is limited (free) tool support for certain knot creation and interaction tasks, which could be useful for lecturers and students within University courses. We present KnotSketch, a tool with multiple functionalities including the ability to: (i) read off a form of Gauss code for a user sketched diagram; (ii) generate a diagram from such a code; (iii) regenerate a knot diagram via a different projection, thereby producing examples of equivalent knot diagrams that may look very different; (iv) interaction capabilities to quickly alter the knot via crossing changes and smooth the curves of the sketched diagram; (v) export facilities to generate svg images of the constructed knots. We evaluate KnotSketch via a case study demonstrating examples of intended usage within an educational setting. Furthermore, we performing a preliminary user study to evaluate the general usability of the tool.
Keywords: *knots, sketching, gauss code, diagram generation.*

1. Introduction

Knots occur within both art and science, and there are many important scientific application domains (e.g. DNA supercoiling [22, 8], quantum wavefunctions [16]). The mathematical field of Knot Theory has been studied extensively, providing a rigorous study of an idealised form of knots (essentially closed loops in 3-space); see [27] for a standard mathematics graduate text book, and [19] for a recent approach aiming to utilise the potential of computers within the field. A standard mathematical approach is to define objects under consideration, provide a formal notion of equivalence, and then to investigate means to try to classify the objects (given two objects, decide if they are equivalent or inequivalent). In this context, two knots are equivalent, called ambient isotopic (isotopic for short), if there is a continuous deformation of the whole of 3-space taking one embedded loop to the other. Knots can be studied thus, making use of knowledge of topological methods.

However, knots can also be studied combinatorially, via knot diagrams (which are regular projections of knots onto a plane, so any crossings are transverse double points). Following Reidemeister [25], isotopy can be realised via diagram transformations: two knots are isotopic if and only if their diagrams (which can be projection onto any suitable plane, so the diagrams can look quite different) differ by a sequence of local diagrammatic transformations (shown later, in Figure 4). Figure 1 shows an instance of the core problem of asking whether diagrams represent the same knot; no answer is provided here, leaving the reader to try to discover an answer for themselves, thereby getting an initial feeling for the challenge via this small example. This question is posed at the beginning of an undergraduate Knot Theory course at the University of Brighton, UK. Since there can be infinitely many diagrams of each knot (considering that different planes of projection can be used, and one could move parts of the knot prior to projection), identifying knot equivalence by comparing diagrams and trying to demonstrate isotopies can be extremely difficult.

Consider the educational context of a knot theory course, with a professor teaching, setting and marking assessments of students, whilst the students are learning, studying, and completing the assignments. The act of drawing of knots can be time consuming and error prone. For example, consider students trying to reproduce the knot shown in Figure 2 in class by hand; this is actually a diagram of the unknot (i.e. it is

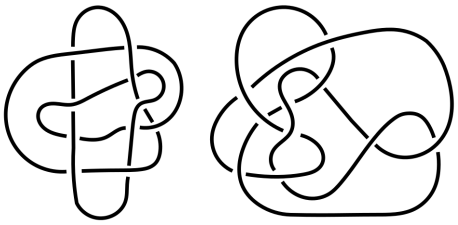


Figure 1: Are these two knots equivalent (isotopic)?

DOI reference number: 10.18293/DMS2016-035

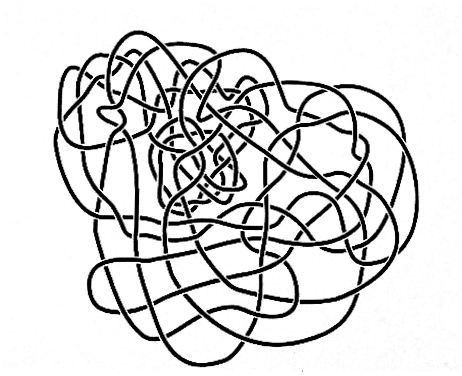


Figure 2: A knot drawing with potential for reproduction errors

equivalent to a knot diagram with no crossings, which can be depicted by a single circle), called the Haken unknot. This motivates the need for a knot drawing tool, with export facilities, and in particular the incorporation of sketch based input facilities to mimic the normal knot drawing method when constructing knots by hand.

Knot codes (like a Gauss code, discussed in Section 2.1) provide a string based representation of the crossings (and their relative orientations) in a knot diagram. In the process of understanding their usage, a student may be asked to produce the code of a given diagram or to try to construct a diagram, given a code, for relatively simple cases. The question "which signed Gauss codes are realisable as (classical) knots?" [17] was open for a long time, but algorithmic solutions have been devised (e.g. [9, 23, 28]). Kauffman [20] extended the class of classical knots (the knots we have described are termed classical knots) to virtual knots, permitting extra crossings without any over/under information, so that every Gauss code is then realisable as a virtual knot. In this paper, we focus on the classical knot case, as would be common in the majority of knot theory courses, but we adopt an underlying approach that can be naturally extended to permit the consideration of virtual knots. The related requirements for a tool are: the automatic reading of the code from a knot drawing, and the automatic generation of a knot diagram from a code.

The codes considered only determine the drawing of a knot diagram drawn on a sphere instead of on a plane; a choice of face (in the shadow of the knot on the sphere, which is the underlying graph given by forgetting the over/under information at the crossings) for stereographic projection is required to provide a drawing in the plane. However, this property can be turned into an interesting benefit. Different choices of (outer) face yield isotopic diagrams (they are, after all just different views of the same knot) but they may look very different. Thus, the alternative choices of outer face give rise to a set of diagrams which are all equivalent but may not look like it, thereby: (i) giving rise to easily constructed examples of equivalence for students to explore; and (ii) helping students to develop their understanding of stereographic projection. Without any tool support, students may find taking a diagram in the plane, considering it as drawn on the sphere and re-projecting using a different outer face, a rather challenging task. It also provides the teacher (or a professional mathematician if the context of intended usage was widened) with a means of examining this somewhat less-familiar relationship between knots that may involve very complex sequences of Reidemiester moves to realize as an isotopy.

We implemented the KnotSketch tool following these elicited requirements. We demonstrate some of its functionalities via a case study, providing a series of worked examples. We follow this up with a preliminary user study to evaluate the general usability of the tool. There are many technicalities to be dealt with in the formal setting of knot theory, but for accessibility to a Computer Science audience we adopted a relatively informal approach, skipping over some technical details (e.g. we presume all knots are tame to rule out pathological examples). An interested reader can refer to graduate level text books on knot theory (e.g. [15, 18]) for more complete details. Discussions about knots (single loops embedded in space) extend to links (disjoint unions of knots) and KnotSketch also supports links.

2. Preliminaries

We describe some facts about knots and their diagrams (informally indicating the content of well known definitions and theorems), and introduce Gauss codes, with examples, providing rationale for the chosen form adopted. A link is a disjoint union of knots, and a link *diagram* is the image of a regular projection (i.e. the only singularities are transverse double points) of the link L with over/under information added at each of the double points, called *crossings*. Every (tame) link has a diagram. Ambient isotopy is an equivalence relation on knots (or links). Each equivalence class of knots is called a *knot type*; equivalent knots have the same knot type. As is common, we abuse language and use the term 'knot' to mean the whole equivalence class (a knot type) or a particular representative. When we say that two knots are different (not equal) we mean that they are inequivalent (i.e they have different knot types). If a knot has the same type as the trivial knot then we say it is *unknotted*. We can orient a knot by nominating one of the two directions along it. If K is an oriented knot then the knot with the opposite orientation, denoted by $-K$, is called the *reverse* of K. The knot $-K^*$ is called the *inverse* of K, where K^* is the mirror image of K (obtainable by switching all crossings in a diagram of K). For oriented links, we can assign to each crossing c of a diagram, its *sign*, which is a value in ± 1, denoted $sign(c)$, as depicted in Figure 3.

Let R_1, R_2 and R_3 denote the diagrammatic moves shown in Figure 4. Two diagrams differ by one of these moves if they are identical outside a small region, and inside the region they differ exactly as shown in the moves. These moves are called the *Reidemeister moves* [25]. We can think of R_3 as moving one of the strands across the crossings of the other two strands. The moves are presumed to preserve orientation. Note that we

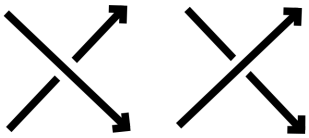


Figure 3: The overpass is the unbroken line for each crossing (on left figure: top left to bottom right; on right figure: bottom left to top right) and the underpass is the broken line. The sign of a crossing is $+$ if traversing along the the underpass, following the orientation, at the crossing the overpass is passing from the left to the right (as in the figure on the left), whilst the sign is $-$ if the overpass is passing from the right to the left (as in the figure on the right).

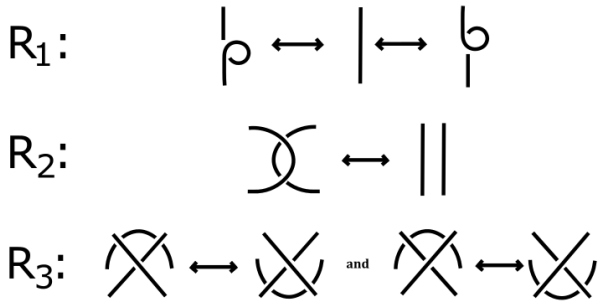


Figure 4: The Reidemeister moves, which are diagrammatic moves encapsulating knot equivalence.

interpret diagrams that differ by homotopy preserving the arc and crossing structure (i.e. the arcs can be moved without changing the underlying crossing structure; e.g. making the knot more "wiggly", or scaling the diagram) as the same.

We say that diagrams D and D' are *isotopic* if D can be obtained from D' by a sequence of moves of type R_i, with $1 \leq i \leq 3$. The diagrams are *regularly isotopic* if R_1 is not used. The following important theorem (see [25]) allows us to study knots and links combinatorially, via diagrams: Suppose that links L_0 and L_1 have diagrams D_0 and D_1, respectively. Then L_0 and L_1 are ambient isotopic if and only if D_0 and D_1 are isotopic.

2.1. Codes

Gauss codes are a means of capturing information in a knot diagram. We adopt a richer form of the codes used in the literature from which others can be recovered. The code of a diagram is given by numbering the crossings, picking an orientation on the curves and then traversing the curves one at a time, writing down the crossings met in order in a complete circuit, noting whether one was on the overpass or underpass (o/u) at each crossing along with an associated sign ($\pm$). It is well defined up to the equivalence relation generated by these choices. Codes for different components of a link are separated by a $/$.

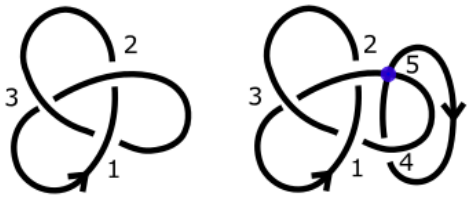


Figure 5: Left: The enhanced Gauss code $o_{1+}\, u_{2-}\, o_{3+}\, u_{1-}\, o_{2+}\, u_{3-}$ indicates the crossings met upon traversal of the knot, together with extra information (over/under, and the sign of the associated immersed curve), realisable as a (classical) knot diagram. Note that the "usual" Gauss code would assign $-$ to every crossing number, yielding: $o_{1-}\, u_{2-}\, o_{3-}\, u_{1-}\, o_{2-}\, u_{3-}$. Right: the code enables the explicit indication of the position of virtual crossings: $o_{1+}\, u_{2-}\, o_{3+}\, u_{1-}\, o_{4-}\, v_{5+}\, o_{2+}\, u_{3-}\, /\, u_{4+}\, v_{5-}$ for the 2-component virtual link. Here, we emphasized the virtual crossing by placing a dot at the crossing (numbered 5).

We describe variations of the code used in the literature. Kauffman, in [20], uses o/u, and the sign of the crossing ($+/-$ described earlier) is attached to each symbol (denoting the number assigned to the crossing), with the same sign occurring for both of the occurrences of this symbol in the code. However, Kurlin [21] includes the same sign of the crossing ($+/-$) but only attaches it to the symbol associated to the undercrossing, thereby removing the need to explicitly indicate o/u's, since the presence of a sign indicates an undercrossing, whilst its absence indicates an overcrossing. Carter [11], considers immersed curves (so the arcs pass through each other at the crossing instead of passing over and under), so there are no o/u's to consider. A variation of the sign convention is adopted, with one $+$ and one $-$ associated to each occurrence of a symbol for a crossing. An intuitive method for calculating this (enhanced) sign convention used is to imagine that the arc under consideration (as we traverse the curve and it passes through a crossing we write down one instance of the symbol for that crossing and decide on its sign) is an undercrossing and calculate the sign as per the earlier method for knots. This leads to one of the two symbols being assigned to a crossing being a $+$ and the other a $-$.

We make use of a variation, which we call the *enhanced Gauss code* (*code* for short) in which we use the sign convention for immersed curves, following Carter, as well as the o/u information (as used by Kauffman) for knots. This generalisation provides us with the ability to deal with curves that do not have any over/under information at the same time as those that do. One application is to explicitly encode virtual crossings in virtual knot diagrams (rather than simply ignoring their presence in the code), permitting the expression of over/under or "through/virtual" at each crossing. The "through/virtual" option can naturally be represented by a v for virtual crossing in the virtual knot diagram setting. Whilst we do not explicitly consider the use of virtual knots within the educational con-

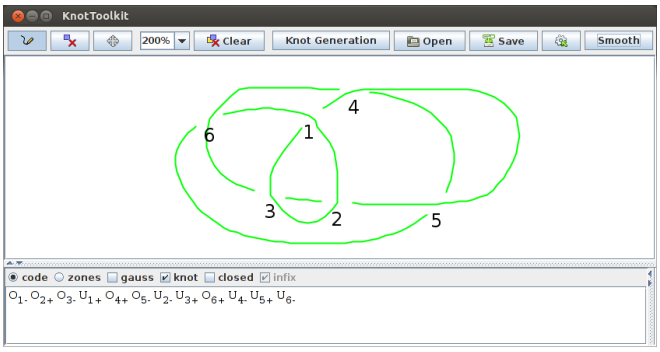


Figure 6: A screenshot of the KnotSketch interface, with *toolbox* at the top, a diagram shown in the *Sketch View* and its automatically produced code in the *Code View*.

text in this paper, we envisage the future usage of this functionality (for advanced postgraduate students or professional mathematicians), and so we decided to adopt a notational convention that would be consistent with this.

Figure 5 shows examples of the (enhanced Gauss) codes of a classical knot diagram (top) and a virtual link diagram (bottom). To enable a comparison with the immersed curve codes, one can consider the shadow of these classical knot diagrams (i.e. make all crossings "through/virtual"); the effect of this on the code is to simply remove the occurrences of o/u, replacing them with v's. This would have necessarily been slightly more complicated if we had used the classical knot sign convention in the code. KnotSketch enables the automatic interpretation of the code of (sketched) link diagrams.

3. KnotSketch

We describe KnotSketch and its main functionalities. KnotSketch is partly based on ink recognition techniques, previously developed for other applications [6, 13, 14].

3.1. User Interface

As shown in Figure 6, the KnotSketch interface is divided into three parts. The upper part is a *toolbox* with buttons to perform the following operations (from left to right):

- change the input mode to *draw*;
- change the input mode to *erase*;
- change the input mode to *move*. At present it is only possible to move whole curves, not parts of a curve (but this would be a useful future additional functionality for interaction);
- change the zoom level;
- clear the current drawing;
- launch the automatic diagram generation facility through a new dialog. The dialog (see Figure 7) contains a text field to enter the code. It is possible to use the clipboard to paste the automatically produced code of a sketched diagram. The code specifies the drawing of the knot on a sphere. By clicking on the OK button, a new dialog is opened (an example is presented in Figure 10 of Section 4). This dialog requires the user to choose a face assigned to be the outside region under stereographic projection from the sphere onto the plane from a list of possible choices (one for each region). A preview pane is shown by the side of the list to assist the user. Moreover, it is possible to personalize the diagram appearance with different visualization options, such as varying the number of spring embedding iterations;
- open a previously saved diagram. If a file describing a diagram drawing (see below) is opened then its content is shown in the *Sketch View* described below, while if a file containing the knot code is opened the diagram generation is launched;
- save the current diagram in a native drawing file format. It is also possible to export a drawing in *svg* format (optionally as an *html* file), whilst the code can be exported in *txt* or *html* format. This facilitates the use of the tool for the production of diagrams that occur in research papers, student submissions, or to include in web-pages;
- open the Options window in order to alter tool settings, such as the size of the gap used in the visualisation of the crossings;
- perform knot beautification using EulerSmooth [30]. Given a drawn diagram, the user can apply a smoothing operation via EulerSmooth to improve the quality of the sketched drawing; EulerSmooth was defined to work with Euler diagrams but it can be heuristically applied to knot diagrams. The smoothing function can be manually controlled and can be applied with various parameter settings.

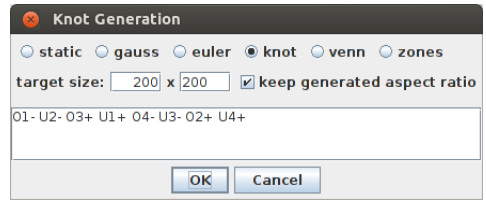


Figure 7: Code to diagram generation box.

In the middle part of the KnotSketch interface the *Sketch View* contains the drawing canvas. The curves are arbitrarily shaped and must be completed with a single pen stroke. Once a stroke has been entered, its endpoints are automatically joined to close the curve. If the user tries to make the curve closed by passing the final part of the curve over the initial part, then any additional crossing created is erased. Upon its completion, a curve is coloured with a colour chosen from a pre-defined list (the predefined colour list can be configured by the user). As soon as a curve is completed, a numeric label is automatically assigned to each crossing point. There are two segments passing through each crossing, and by default the last drawn segment is shown as an over-crossing (as a future option, the user will be able to choose the default behaviour, permitting the automatic provision of alternating

crossings as the knot is traversed, for example); the crossing type (over/under) can be manually changed by the user by clicking on the crossings. The length of the gap used to indicate the crossing can also be configured by the user.

In the lower part of the interface, the *Code View* contains the automatically generated enhanced Gauss code, as described in Section 2.1 (it is possible to select the code and copy it to the clipboard). At the top of this view a toolbox provides a set of options through which the user can select some other forms of codes (that are not required for this particular application).

A user study is discussed in Section 5, and Figures 8-12 show some related screenshots of the tool.

3.2. Back-End

KnotSketch is written as a Java 7 application. Other than knot drawing, the main features of the application are knot interpretation (i.e. computing the code of a drawn knot), knot generation from code and knot beautification. The enhanced Gauss code, displayed in the *Code View*, is incrementally constructed and is updated every time a curve is added or deleted. It is stored in an internal format enabling efficient operation executions. A knot is represented as a closed polyline.

KnotSketch can generate knot drawings from an enhanced Gauss code using a planar graph based construction: the graph contains one node for each crossing, and its combinatorial embedding of the graph can be calculated by using the crossing signs $(+/-)$ and the choice of face assigned to be the outer region. Then an algorithm from the OGDF framework [24] can be used to embed the planar graph in the plane. The user can optionally modify the embedding by selecting a number of iterations of a force directed algorithm [5] to apply. The algorithm theoretically preserves the regions, but approximation due to rounding could cause non-preservation. However, the tool provides a post-check if the initial and final diagram codes are equivalent. Finally, by traversing the graph edges appropriately, the curve for the knot diagram is constructed.

The knot beatification step can be performed by using the Ocotillo java library from EulerSmooth. The drawn knot is treated as if it was an Euler diagram (essentially taking the shadow of the knot, viewing all crossings as if they passed through instead of over/under each other; note that the constraints of non-self intersections for Euler diagrams must be relaxed in this context). The knot diagram is converted to the format used by Ocotillo and the beautification process is performed on it, using the parameters set by the user. After each beautification iteration the diagram is converted back to the original format in order to display the beautification progress to the user. Future enhancements of EulerSmooth in this context should lead to enhancements of the outputs.

4. Case study

We describe a set of activities that can be performed with KnotSketch within an educational context, demonstrating the capabilities of the approach and the tool. We highlight a novel viewpoint that permits an uncommon form of user exploration, which would be challenging for students to perform without tool-based assistance. We discuss general activities that users may undertake along with low level tasks that users may perform along the way.

Consider the task of starting with a code and creating a diagram of a knot with that code. This task can be used to test a users' understanding of the code to diagram process, and to highlight how challenging the construction can be without further assistance. In a simple case, the user can directly sketch the requested diagram. They can check their solution via the automatic production of the code, displayable in the *Code View*. If the knot produced differs from the required solution via crossing changes alone (i.e. switching over and under crossings) then user interaction via clicking on the crossings can be applied to alter the diagram accordingly. If the diagram drawn has the wrong code and the code differs by more than crossing changes then the curve needs to be redrawn. An alternative approach is to use the automatic diagram generation facilities from a code (described in Section 3). For the task of producing a diagram with the correct code, any choice of outer face will do, so the user can select any of the proposed faces. This functionality leads to an interesting possibility for the exploration of knot equivalence (isotopy). Given a code one can generate different views of the knot (one for each face, by using the different choices of outer face in the plane of projection). Since each of the views are diagrams of the same knot viewed from a different perspective, all of these diagrams are equivalent. However, they may look very different and it may be challenging for a user to construct an isotopy (a sequence of Reidemeister moves) from one diagram to the other. Thus, given two knot diagrams that potentially look very different, but which are in fact different projections of the same knot, the user can try to identify this knot equivalence by searching through the options and visually comparing the outputs.

Figure 8-12 show diagrams from the tool similar to those of the case study. In detail, Figure 8 shows a user sketched (and smoothed) drawing of a trefoil knot with 3 crossings, its automatically produced code, along with a more complicated knot with 7 crossings. Figure 9 shows diagrams of a knot K, the knot K' with a single crossing change, and, $m(K)$, the mirror of the knot K, which is shown as K with all crossings changed; the effects on the code are also shown. Figures 7 and 10 show an example of a code being used to generate a knot, demonstrating the user's view at the time of selection of outer face. The code-like information displayed for each choice is a means of specifying the region (details are omitted since they are unimportant for our purposes here); we display drawings arising from all six choices of outer face for this code. Figure 11 shows diagrams (far left and far right) and their (not obvious) equivalence as a sequence of Reidemeister moves (indicated by R_i's). These two diagrams can be seen to be equivalent via projection onto different faces; the left and right diagrams are those shown in Figure 10(f) and Figure 10(e), respectively, up to moving the arcs without changing

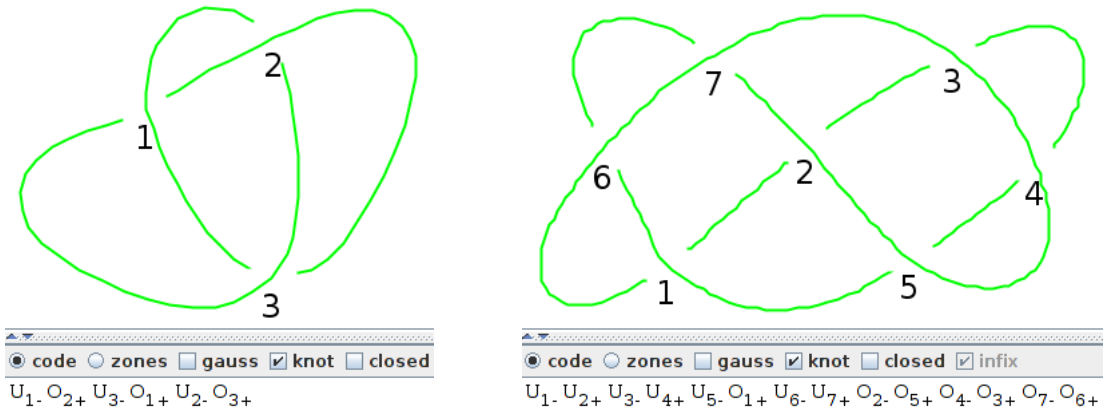


Figure 8: Screenshots of a trefoil knot (left) and a knot with seven crossings (right) and their codes.

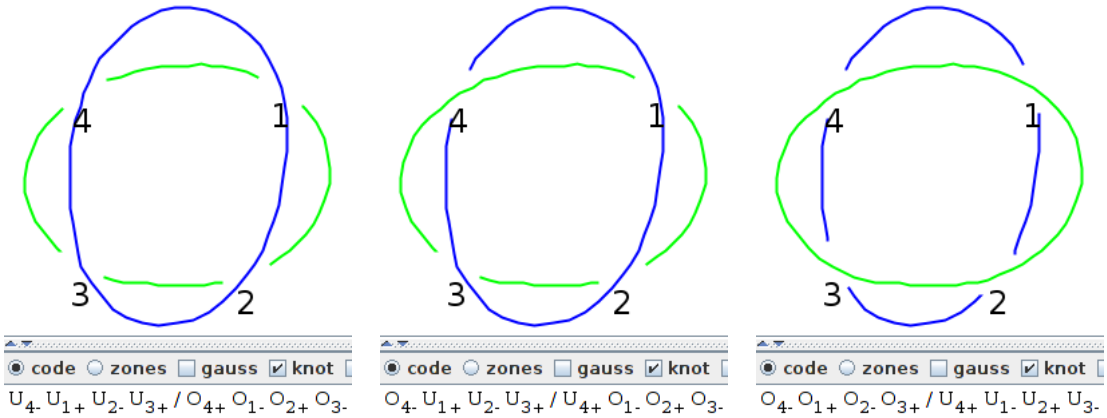


Figure 9: A diagram of a link (left), the diagram obtained by changing crossing 4 (middle), and the mirror of original diagram, given by changing all crossings (right).

the crossing structure. Figure 12 shows an example of generation of a knot diagram from a complicated code, showing the initial creation of a knot, along with an initial improvement using the EulerSmooth functionalities.

5. Experimental setup

We performed a preliminary evaluation on a student population to test the usability of the tool itself. In particular, we measured the perceived usability of the system through a System Usability Scale (SUS) [7] questionnaire. The questionnaire is composed of 10 statements to which participants assign a score indicating their strength of agreement on a 5-point scale. The final SUS score ranges from 0 to 100. Higher scores indicate better perceived usability. We also gathered some participants' free-form comments during and after the experiment.

The demographic mostly consisted of Computer Science students, which may help with familiarity with the use of software. In the future, we will also examine the usage with Mathematics students whom may be more familiar with the mathematical context, but perhaps less proficient in software usage. We recruited a total of 10 participants (3 female). The ages ranged from 21 to 51 (with mean $M = 31$, and standard deviation $SD = 10.4$). All of them were habitual computer users and had previous experience with touch-screens. None of them had any prior knowledge of knot theory.

The set of tasks we considered were as follows.

TASK 1 Given the following knot code, generate a diagram of it using the "generation from knot code" functionality of KnotSketch: $o_{1+}\ o_{2-}\ u_{3+}\ u_{1-}\ u_{4+}\ u_{5-}\ o_{5+}\ o_{4-}\ u_{2+}\ o_{3-}$;

TASK 2 Given the knot drawn on the sheet, reproduce it using KnotSketch and save its knot code in a text file. Moreover, save the drawing in the tool's native file format (this task was repeated three times - TASK 2.1, TASK 2.2, TASK 2.3 - once for each knot in Figure 13);

TASK 3 Open the native format of the drawing previously produced in TASK 2.2 and generate its projection onto a

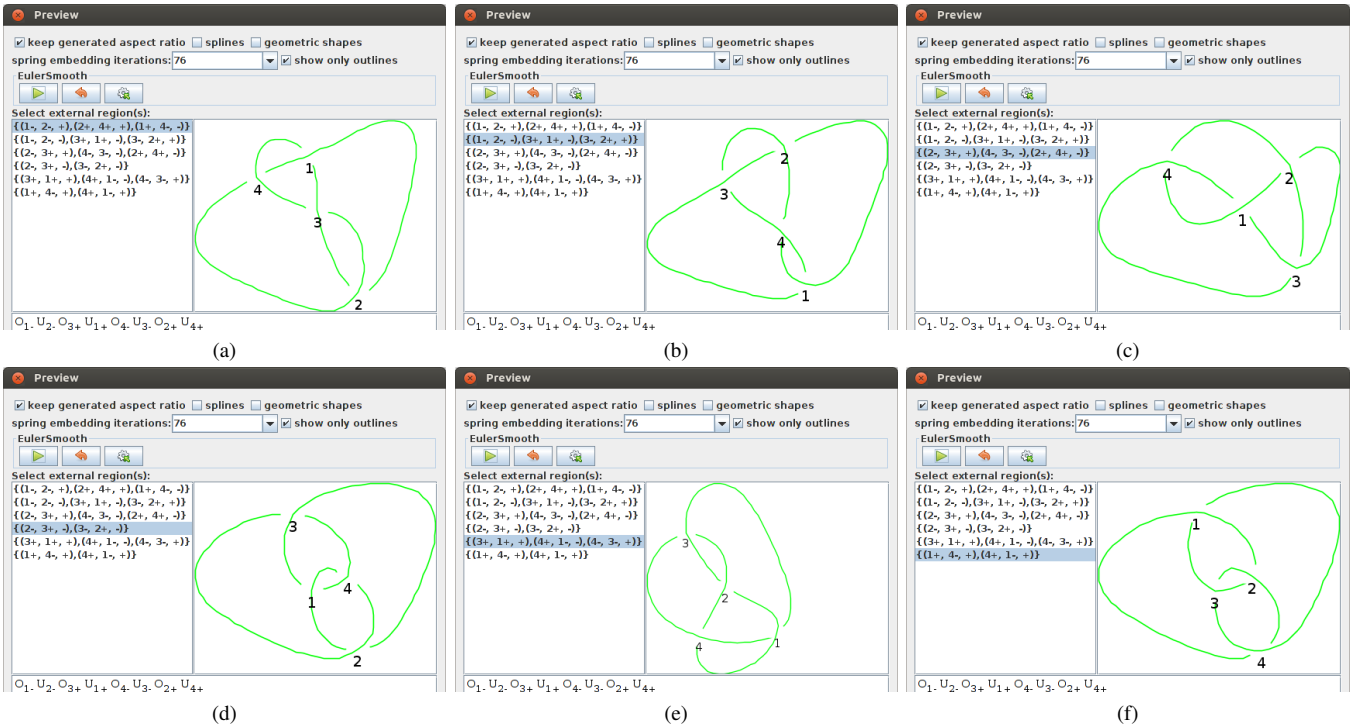


Figure 10: Displaying the preview of the different choices of outer face for the projection, using the code from Figure 7.

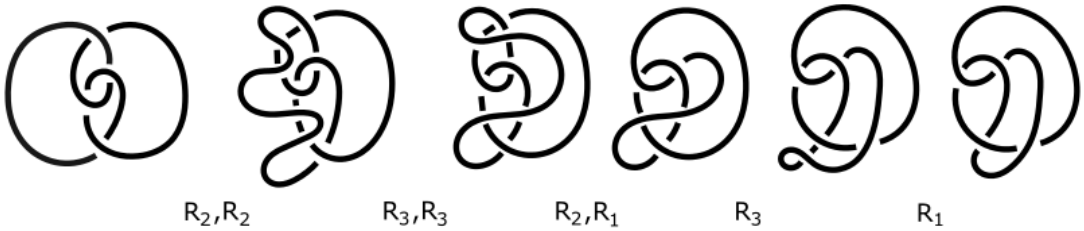


Figure 11: A sequence of Reidemeister moves demonstrating the equivalence between the left diagram (c.f. the diagram in Figure 10(f)) and the right diagram (c.f. the diagram in Figure 10(e)).

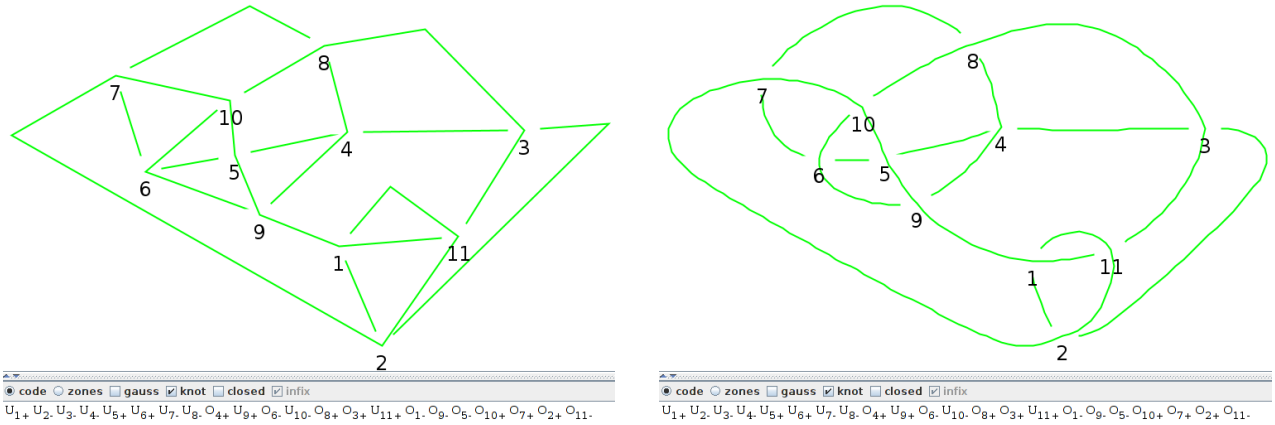


Figure 12: Generation of a knot diagram from a complicated code, with the initial creation of a knot (left), along with an initial improvement using the Eulersmooth functionalities (right).

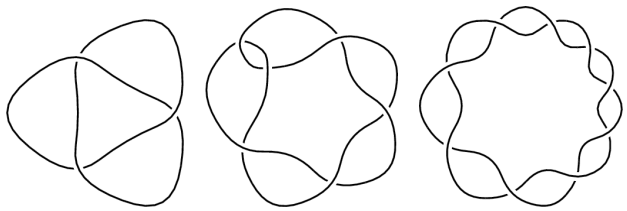
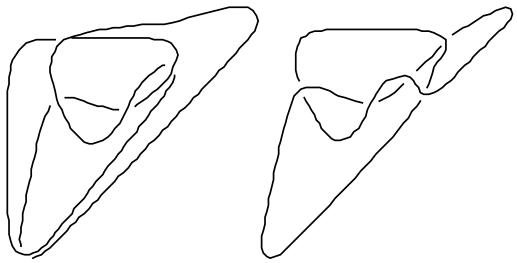
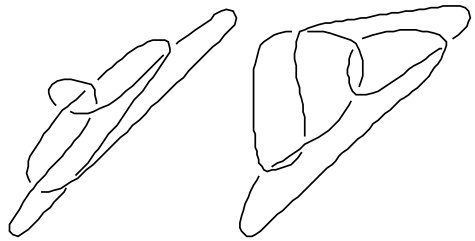


Figure 13: TASK 2 knots.



(a) Pair of knots for TASK 4.1.



(b) Pair of knots for TASK 4.2.

Figure 14: TASK 4 knot pairs.

different outer face;

TASK 4 Given the pairs of knot diagrams on the sheet, indicate if they differ by redrawing via projection onto a different outer face or not (the task was repeated twice - TASK 4.1, TASK 4.2 - once for each knot pair in Figure 14).

The time limit for the completion of each subtask was 5 minutes.

The tasks were executed on a *Dell Precision T5400* workstation equipped with an *Intel Xeon* CPU at 2.50 GHz running *Microsoft Windows* operating system and the Java Run-Time Environment 8. The device used for the experiment was a Sympodium ID250 Interactive Pen Display, attached through both USB and RGB cables to the work station.

5.1. Results

The tasks have been successfully completed by almost all participants. Only two tasks were not completed successfully by all participants: TASK 2.3 and TASK 4.1, with two errors each. The mean task completion times are reported, along with standard deviations, in Table 1. As expected, the subtasks of TASK 2 showed increasing difficulties; in particular for TASK 2.3, some users needed to attempt the drawing several times. However, the error on this task was simply an incorrect sign on a single crossing. TASK 4 presented some difficulties for most participants, with relatively high execution times and two errors for TASK 4.1.

Task	Avg time (sec.)	S.D.	User errors
TASK 1	64.7	16.6	0
TASK 2.1	45.3	14.1	0
TASK 2.2	60.2	13.3	0
TASK 2.3	137.6	72.8	2
TASK 3	20.0	11.8	0
TASK 4.1	135.0	48.8	2
TASK 4.2	152.6	41.6	0

Table 1: Task completion times.

Username	Score
participant 1	77.5
participant 2	85
participant 3	62.5
participant 4	75
participant 5	87.5
participant 6	75
participant 7	85
participant 8	77.5
participant 9	92.5
participant 10	85
Average	80.25

Table 2: SUS-like questionnaire scores for user satisfaction.

The scores of the questionnaire calculated from the responses of the participants are shown in Table 2, indicating scores ranging from 75 to 92.5, with an average value of 80.25. This value indicates a good level of satisfaction [4]. An analysis of the questionnaire scores showed that the participants judged the tool very useful for its purpose, but they did not find all the features simple to use and some participants needed some learning time to be able to master them. For instance, some participant lamented the lack of feedback on the presence of errors in a written code; others had problems in finding out how to generate the projection of a drawing onto a different outer face.

Similar considerations emerged after a brief interview/discussion with the participants involved in the experiment to understand what aspects of the tool were of help in the execution of the tasks: probably due to unfamiliarity with Knot Theory, some participants expressed initial doubts about the comprehension of the tool's purpose. Furthermore, some of them had difficulty in completing TASK 4 even with the help of the tool, with one participant stating: "I would have preferred more support from the tool to perform TASK 4". Nevertheless, he declared "I would never be able to complete that task without the help of the tool".

Tool	KnotSketch	KnotPlot [26]	Knot ID [2]
Hand sketch	Yes	Yes (but requires license)	No (point and click/drag)
3D visualisation	No	Yes	Yes
Link diagrams	Yes	Yes	No
Save/Load/Export diagram	• Supports saving and loading diagrams • Diagrams can be exported in SVG and HTML format	• Supports saving and loading diagrams • Diagrams can be exported in EPS format	• Load list of coordinates • Load gauss code (no visualisation of the knot in this case)
Edit diagram	• Supports adding, deleting and moving individual components	• Supports adding, deleting and moving individual components or a segment of any component	• Supports adding, deleting or modifying individual segments of any component
Diagram generation via code	Supports Gauss Code	Supports Dowker code and Braid word	No generation (but identification available from Gauss Code)
Diagram reprojection via outer face	Yes	No (but rotation of 3D visualisation is possible)	No
Diagram smoothing	Yes	Yes	No

Table 3: Visualisation features of KnotSketch, KnotPlot and KnotID

6. Related Work

Existing knot tools have varied functionalities. Table 3 provides comparison of visualisation aspects of KnotSketch with KnotPlot [26] and Knot ID [2]. We provide some further details and briefly discuss other notable tools below. To the best of our knowledge, the re-projection facility onto different outer faces offered by KnotSketch is novel, as is the capability to deal with (sketching, interpreting) virtual knots or links (note that the study was restricted to classical knots and links).

KnotPlot [26] is a widely used program for visualising and interacting with knots that has theoretical underpinnings as described in [29]. A knot can be sketched by hand (in the upgraded version requiring a license) or constructed from code using a tangle notation system developed by Conway [12], or from a braid word, or a Dowker-Thistlethwaite code (which does not uniquely specify composite knots [26, page 98]). A knot can be refined into a smoother configuration using a 3D technique, which is computationally expensive, especially for large knot diagrams with more than 50 crossings [26, page 105]. KnotPlot has a built-in database containing a wide range of knots and links which can be viewed, modified and saved, along with other features such as: computing the Alexander and HOMFLY polynomials, the writhe, the average crossing number, the thickness and the Dowker code of a knot; searching for minimal stick conformations and interesting random knots; generation of arbitrary braids; enabling the consideration of open knots or links; interactive construction of knotted and linked spheres in four dimensions.

KnotID [2] is a web application that allows the viewing of topological information about knots. The application can be used: (1) to import three-dimensional curves (as a list of 3D coordinates or by generating torus knots from two coprime numbers); (2) to draw two-dimensional curves; (3) to directly input an enhanced version of the Gauss code similar to those described in Section 2.1. In addition to displaying topological information, the tool can, through the application of topological invariants, identify if the input knot is equivalent to knots found in a lockup table (based on [1]). It permits computation of properties such as the Reduced crossing number, Determinant $|\Delta(-1)|$, $|\Delta(\exp(2\pi i/3)|$, $|\Delta(i)|$, and Vassilev invariants. Compared to KnotSketch, the software offers a non-sketching oriented drawing interface and no features such as delete/smoothing/load/save or the ability to display the crossing numbers. Moreover it does not offer functions to generate a knot from a gauss code or to regenerate a knot by re-projection onto a different plane.

Libbiarc [10] can be used to manipulate and analyze properties of knotted curves, compiled into a $C++$ library; the most frequently used functions are available as an API. The core library provides functions to interpolate point-tangent data with bi-arcs, access information such as curvature and torsion, compute the length and thickness. The library includes a viewer application, called curview, for visualisation, in which a knot can be loaded, manipulated and saved.

KnotApp [31] is a thesis that describes a program that displays a knot and other objects, such as the knot's crossing map and its trisecant curve, using the jReality JOGL Viewer. The knot may be chosen from a provided list or loaded from the file system. Then the chosen knot can be edited by dragging its vertices using the provided knot editor. The crossing map implemented displays the set of irregular projections of the underlying knot as curves on the unit sphere. The set of trisecants of the underlying knot is visualised as a curve in the 3-dimensional torus. However, the KnotApp application itself does not appear to be currently available to be tested.

Other notable knot tools, with varying functionalities include Linknot (see [19]), Knotscape [32] (uses Dowker-Thistlethwaite codes), and MING [33] (which minimizes MD-energy of polygonal knots, reduces the numbers of edges, and draws/visualizes their 3-dimensional pictures).

7. Conclusions

We have developed KnotSketch, a tool to facilitate knot exploration, with particularly interesting functionalities of permitting knot sketching followed by classical knot diagram regeneration via re-projection onto different planes via stereographic projection from the diagram drawn on the sphere. This could be particularly useful in University level educational settings, and we examine the potential capabilities via a case study. In addition, we adopt an enhanced version of Gauss code that enables the future handling of virtual knots. This ability to deal with virtual knots would likely be useful in the context of expert users, such as professional mathematicians making use of the tool in exploratory mathematics or simply to produce figures for use in publications (the use of classical knots is of course pertinent to this user class as well). One can also envisage different semantical interpretations of these types of diagrams with mixed crossing types, and this would widen the applicability even further beyond the consideration of knots.

We performed a preliminary user-based empirical study to gain insight into the usability of the toolkit, with an aim for future improvements. We plan to perform more extensive user testing, and adopt an iterative developmental methodology taking into account user-insights after testing periods. We wish to explore the utility for both Mathematics students and professional Mathematicians. Using the interface at KnotInfo [3], a user can select sets of knots and request information about them including their diagrams or compute many invariants. This interface and database (and others) could be made to be usable in conjunction with KnotSketch, thereby greatly enhancing its functionalities.

References

[1] Knot atlas take home database. At `http://katlas.org/wiki/The_Take_Home_Database`.

[2] Knot ID website. At `http://inclem.net/knotidentifier/`.

[3] Knot info website. At `http://www.indiana.edu/~knotinfo/`.

[4] A. Bangor, P. T. Kortum, and J. T. Miller. An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, 24(6):574–594, 2008.

[5] F. Bertault. A force-directed algorithm that preserves edge-crossing properties. In *Graph Drawing*, volume 1731 of *Lecture Notes in Computer Science*, pages 351–358, 1999.

[6] P. Bottoni, G. Costagliola, M. De Rosa, A. Fish, and V. Fuccella. Euler diagram codes: Interpretation and generation. In *Proceedings of the 6th International Symposium on Visual Information Communication and Interaction*, VINCI '13, pages 105–106, New York, NY, USA, 2013. ACM.

[7] J. Brooke. Sus: A quick and dirty usability scale. In P. W. Jordan, B. Weerdmeester, A. Thomas, and I. L. Mclelland, editors, *Usability evaluation in industry*. Taylor and Francis, London, 1996.

[8] D. D. Buck. EP/H031367/1 SANDPIT : Knots and Evolution - Topologically Driven Integrase Mutagenesis, and EP/G039585/1 DNA Knotting and Linking: Applications of 3-Manifold Topology to DNA-Protein Interactions, 2009.

[9] G. Cairns and D. Elton. The Planarity Problem II. *Journal of Knot Theory and its Ramifications*, 5:137–144, 1996.

[10] M. Carlen. *Computation and Visualisation of Ideal Knot Shapes*. Ph.d. thesis, EPFL, Lausanne, 2010.

[11] J. Carter. Classifying immersed curves. In *Proc. Amer. Math. Soc*, volume 111, pages 281–287, 1991.

[12] J. H. Conway. An enumeration of knots and links, and some of their algebraic properties. *Computational Problems in Abstract Algebra*, pages 329–358, 1970.

[13] G. Costagliola, M. De Rosa, A. Fish, V. Fuccella, and R. Saleh. Curve-based diagram specification and construction. In *Proc. VL/HCC*, pages 39–42, 2013.

[14] G. Costagliola, V. Fuccella, and M. Di Capua. Interpretation of strokes in radial menus: The case of the keyscretch text entry method. *Journal of Visual Languages & Computing*, 24(4):234 – 247, 2013.

[15] P. Cromwell. *Knots and Links*. Cambridge University Press, 2004.

[16] M. Dennis and A. Taylor. Vortex knots in tangled quantum eigenfunctions. *Nature Communications*, (12346), 2016.

[17] C. Gauss. Werke. *Band 8. Teubner*, 1900.

[18] N. Gilbert and T. Porter. *Knots and Surfaces*. Oxford Science Publications, 1997.

[19] S. Jablan and R. Sazdanovi. *LinKnot: Knot Theory by Computer*, volume 21 of *Series on Knots and Everything*. World Scientific, 2007.

[20] L. Kauffman. Virtual knot theory. *European Journal of Combinatorics*, 20:663–691, 1999.

[21] Kurlin. Gauss paragraphs of classical links and a characterization of virtual link groups. *Math. Proc. Camb. Phil. Soc.*, 145:129–140, 2008.

[22] R. Metzler. Knots, bubbles, unwinding, and breathing: Probing the topology of DNA and other biomolecules. *Handbook of Theoretical and Computational Nanotechnology, edited by M. Rieth and W. Schommers*, 1:1–54, 2005.

[23] J. Nagy. Über ein topologisches Problem von Gauss. *Mathematische Zeitschrift*, 26(1):579–592, 1927.

[24] OGDF. Open graph drawing framework. `http://www.ogdf.net`.

[25] K. Reidemeister. Elementare begrøundung der knotentheorie. *Abh. Math. Sem. Univ. Hamburg*, 5:24–32, 1926.

[26] R.G.Scharein. *Interactive topological drawing*. Ph.d. thesis, The University of British Columbia, 1998.

[27] D. Rolfsen. *Knots and Links.* AMS Chelsea Press, 2003.

[28] P. Rosenstiehl and R. Tarjan. Gauss codes, planar Hamiltonian graphs, and stack-sortable permutations. *Journal of algorithms*, 5(3):375–390, 1984.

[29] R. Scharein. *Interactive Topological Drawing*. PhD thesis, University of British Columbia, 1998.

[30] P. Simonetto, D. Archambault, and C. Scheidegger. A simple approach for boundary improvement of Euler diagrams. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):678–687, 2016.

[31] M. Sommer. *Visualization in Geometric Knot Theory - Understanding the mathematical structure of trisecants*. Ph.d. thesis, Technische Universität Berlin, 2007.

[32] M. Thistlethwaite. Knotscape. At `http://www.math.utk.edu/~morwen/`.

[33] Y.-Q. Wu. Ming. At `http://www.math.uiowa.edu/~wu/`.

Visualizing student engagement in e-learning environment

Enrica Pesare, Teresa Roselli, Veronica Rossano
Dipartimento di Informatica
Università degli Studi di Bari Aldo Moro
Bari, Italy
{enrica.pesare, teresa.roselli. veronica.rossano}@uniba.it

***Abstract*— The learning assessment in e-learning contexts is one of the latest challenges for educational technology researchers. One of the main issues to be addressed is the definition of dimensions that should be used to measure the learning effectiveness. In this perspective, the research work aims at defining the engagement indicators useful to assess the active participation of students in social learning environments. Moreover, the paper presents the design and implementation of Learning Dashboards aimed at visualizing the student engagement in online communities where the engagement and involvement of students are the key factors for successful learning.**

Keywords- Assessment; Learning Analytics; Learning dashboard; Social learning environments; Engagement.

I. INTRODUCTION

The learning assessment is one of the main issues in the educational field. In the latest years, the interest in this topic is rising thanks to the evolution of e-learning methods and techniques. The new teaching models adopted in the MOOCs (Massive Open Online Courses) [1] [2] require the adoption of new tools and measures that allow teachers to make effective and reliable assessment even with a large population of learners. In the literature, different solutions have been proposed, such as the adoption of qualitative measures [3] [4] to have detailed information about the interaction of the learner with the peers and the didactic resources, or the definition of tools able to supply detailed information on the student learning path, starting from the tracking data of e-learning platform [5] [6]. For these reasons, the interest of the educational technology researchers is focusing on Learning Dashboards that allow, both faculty and students, to tune their actions (lectures, assignments, study) based on rapid feedbacks on student progress in learning [7]. Currently, the interactions are becoming even more complex in the e-learning environments. Thus, the number of dimensions to take into account to make a successful assessment of the student's learning path is improving. The main challenge of Educational Technology field is to point out all the dimensions useful to measure the effectiveness of the e-learning paths [8].

DOI reference number: 10.18293/DMS2016-028

In this perspective, our point of view is that in the e-learning environments 2.0, where the construction of the knowledge is shared among the different users, to measure the engagement in the e-learning process could give information about the learning effectiveness. For this reason, the paper presents a Learning Dashboard aimed at visualizing the student engagement in web based learning environment. In particular, the attention is focused on online communities, where the engagement and involvement of students is the key factor for a successful learning.

II. RELATED WORKS

The Learning Dashboards are commonly used in a wide range of e-learning environments: Learning Management Systems, Web-based and Personal Learning Environments, Massive Open Online Courses and so on. Information Visualization techniques, indeed, are powerful tools in the learning analytics research field, since they allow to visualize the collected data about student's activities. Visualization can impact on the user behavior and motivation, for both students and teachers, and promote self-awareness and reflection about the learning process [9].

The interest in this research field and the number of proposed approaches are growing in the latest years. For ours research goals, we decided to narrow the research on two important aspects in e-learning environments: student's progresses and performances and student's engagement.

A. Visualizing progresses and performaces in e-learning enviroment

Numerous are the solutions proposed to face this issue. Student Activity Meter (SAM) [10] [11], for example, was designed to explore the classroom activities. It provides an overview of student's activities through simple statistics. For each student, the indicators are the time spent and the documents used. These indicators are compared with minimum, maximum and average values of the whole class.

The first indicator, time spent, is displayed with a line chart visualization; it *shows a line for every student in the course; the horizontal axis shows the dates and the vertical axis displays the total time spent* [11]. Indeed the visualization of

activities over time is a key aspect when analyzing student behaviors. The second indicator, document used, is not directly displayed but it is used to provide recommendations with a simple recommender system.

Furthermore a parallel coordinates chart shows correlations among *(i) the total time spent on the course, (ii) the average time spent on a document, (iii) the number of documents used and (iv) the average time of the day that the students work* [11] allowing the discovery of patterns in the student's behavior.

Mastery Grids introduces the so called "social progress visualization" [12] to engage students and to guide them through the learning resources exploration. The indicators used are the progress made by each student. The progresses are presented in a matrix where for each domain topic (horizontal dimension) and kind of resources available for that topic (vertical dimension) the level of completeness *or level of progress a learner has for the given combination of topic and resource type* is showed through color intensity (third dimension).

In addition to the individual progress visualization, Mastery Grids enables comparison with other peers (the class average, the top ranking students) and highlights the differences between the individual user and the group. Furthermore, it enables more compact and detailed comparison selecting one kind of resource. This direct comparison pushes the user to improve her/his activity by stimulating them to complete different kind of activities and to access new content at the same time.

CAM Dashboard was designed to allow the exploration of learners' behavior in Personal Learning Environments and to enable both self-reflection and comparison with peers to improve student's motivation [13] [14]. It collects data from different data sources even outside the traditional LMS, and provides visualization according to the student's goals.

The indicators are the total number of activities done by the student and the number of events or time spent for each application used by the student over time. Furthermore those indicators are grouped by day of the week, by action performed or by resources involved. Noteworthy is the distinction of different kind of actions that enables a deeper level of analysis, for instance the distinction of active and passive actions such as writing and reading activities.

The main visualization is a line chart of the activity over time (annotated timeline) while the grouped indicators are presented with bar chart visualizations that allow the comparison between the time spent by the user and the average time of peers.

VeeU was designed to enhance student assessment in distance learning environment for both teachers and students. [5]. The indicators proposed are the number of daily accesses, the number and distribution of activities in a course and the completion rate of course activities.

The accesses are presented in a time line visualization at different aggregation levels (different courses or single course) for the teacher, while the student visualizes the number of her/his daily accesses compared to the average values for the class. A pie-chart visualization is used for the percentage and distribution of activities while the completion rate is displayed in a gauge chart visualization with a list of recommendations to help student in achieving goals.

Even StepUp! has been designed for students empowerment in open learning environments [15][16][17] in the "quantified self" perspective [18]. It collects tracking data from group blogs and twitter (post, comments, tweet). Even in this case the indicators are the number of activities and the time spent but the activities are also classified in assimilative (blogging and writing reports), communicative (twitter and comments) and productive activities (programming) [19] enabling distinction of active and passive behavior in this case too.

The indicators are presented with numerical data and sparklines for every student to provide a quick overview. The sparklines can be detailed in a stacked bar chart visualization that displays activity over the weeks, grouped by kind of activities and participation to promote awareness of what students did and how they spent their time.

TrAVIS [20] collects data about communication activities in distance learning environments to promote self-monitoring. It distinguishes four levels of interaction (aggregation, discussion, cooperation and collaboration) that correspond to four levels of indicators:

- aggregation level: connection frequency, threads started, messages posted, messages replied, and messages quoted. These indicators are commonly used to describe the activities of each individual student;
- discussion level: browsing, forums, posting, reading and chatting activities. These indicators are useful to estimate the user interest in the discussions or her/his level of community interaction;
- cooperation level: thread started, new messages, replied messages, quoted messages, files uploaded, files downloaded and participation level. These indicators are useful to evaluate the contribution of each user to the community.
- collaboration level: even these indicators are constructed from the lower levels indicator but in this case the focus is moved from the individual perspective to the group level perspective. As a matter of fact they are collected at group level within a defined timespan in order to compare the participation rates or the productivity rates of different groups.

The visualization technique at each level is the spider chart: a spider chart for each user (at the aggregation, discussion and cooperation levels) or for each group (at the collaboration level) allows visual comparison among students and among groups.

B. *Visualizing engagement in e-learning enviroments*

In the e-learning field, the increasing interest in the engagement dimension has led to the definition of different dashboards to monitor and improve student engagement.

VisEN, for example, is addressed to students to allow the exploration of data about engagement in form of visual narratives [21] [22]. It presents the student's engagement in a quick gauge chart while further details are provided on demand. Students can interact with visualizations to realize how engagement score was calculated, based on course interaction (page click), study time (reading and reviewing durations), submissions and questionnaire scores. They can compare their engagement with peers (all class members or similar engaging peers) at different levels (global score or activity level).

Even in the emerging field of MOOCs, student engagement has acquired a great importance. Coffrin [23] proposes the analysis of student's activities in MOOCs using visualization techniques. In this case, the visualizations are used at macro level to understand patterns of student engagement. However, even if the proposed analysis are not strictly related to our goal, they give us some interesting points of view on the topic. The analysis starts from two simple histograms of students' participation and assessment performance, broken down by week. This first visualization confirms the common pattern of high interest at the beginning with increasing attrition rate over time, but breaking down the data, a deeper understanding of engagement is provided. In particular, students in a MOOC can be classified into Auditors, Active and Qualified students according to the kind of activities performed (video lectures and assessments). This breakdown can be useful in different visualization (histogram or state transition diagram) to closely analyze the percentage and relative proportion of students and to understand the temporal evolution of engagement trajectories [24].

Apart from individual participation in MOOC, even the social interactions have a great influence in particular when involving social learning processes [25].

Schreurs [26] proposes the application of Social Learning Analytics in a MOOC platform. In particular, it focuses on the network visualization from discussion forums: every user is a node, every reply to a post is a tie, while a simple tag cloud allows the filtering of ties and nodes based on the contents. Moreover a tooltip will allow the discovery of topic of interest and expertise of every user. Then different network perspectives can promote reflection on learners' interaction or make visible the contents developed through discussions.

Another perspective on engagement in web-based learning environment is given in [27] where the authors propose the visual representation of cognitive and behavioral indicators of engagement to support teachers in monitoring learners' engagement. To reflect the multidimensionality of engagement the indicators are built from student participation actions (behavioral indicators) and from modification actions on the learning documents structures (cognitive indicators). The first ones are computed from number of actions and duration (number of login, number accesses to a learning resource, time spent on a resource etc.) while the latter are computed from a particular subset of action related to the structural modification of learning documents (create, add, update, delete, move and insert).

The proposed visualization is based on the small multiple visualization technique for each student's session: each frame represents a simplified mind map at different times (t_i) with dotted lines for deleted elements and solid lines for existing ones. The small multiple representation can be combined with a linear representation for each node where the structural modifications on the selected node are made visible over time. The small multiple visualization might have problems in case of growing numbers of nodes or growing number of time intervals while the linear representation is easily manageable for temporal data even if it does not allow comparisons of data about different nodes at the same time.

III. VEEU 2.0

The main goal of our research is to make students and teachers aware of their engagement in a social learning environment. As a matter of fact, teachers and students need to be aware of what kinds of interactions are occurring in the virtual space and how the building up knowledge process happens. This is the so called "situational awareness" that, according to Few's principle design [28], is one of the main purposes of dashboards. In this perspective, the rapid perception of information through the dashboard is fundamental to facilitate the decision-making process.

To achieve these goals, in the following sections are briefly presented the social learning environment and the dashboard design process, from the data analysis (to define the most suitable predictors and indicators), to the selection of the best visualization techniques (to depict the relevant data at a glance).

A. *Social Learning Environment*

The learning environment, as depicted in Fig. 1, is a customized Moodle instance, in order to preserve consistency with the learning management system in use in our University. To enable Moodle to adopt the social paradigm the SocialWall plug-in, (https://moodle.org/plugins/format_socialwall) was used. It turns the traditional Moodle course format into a social interface. Moreover, a Wiki activity as collaborative knowledge repository has been added. In particular, the Wiki enables users to co-create a complex web document, even without any knowledge about HTML, allowing them to gather pieces of knowledge and build a shared repository.

The SocialWall enables users to post messages, documents, links and any other kind of resources already available in Moodle. Users can express their like or comment the posts. All the social activities will appear on a timeline provided with some filtering tools.

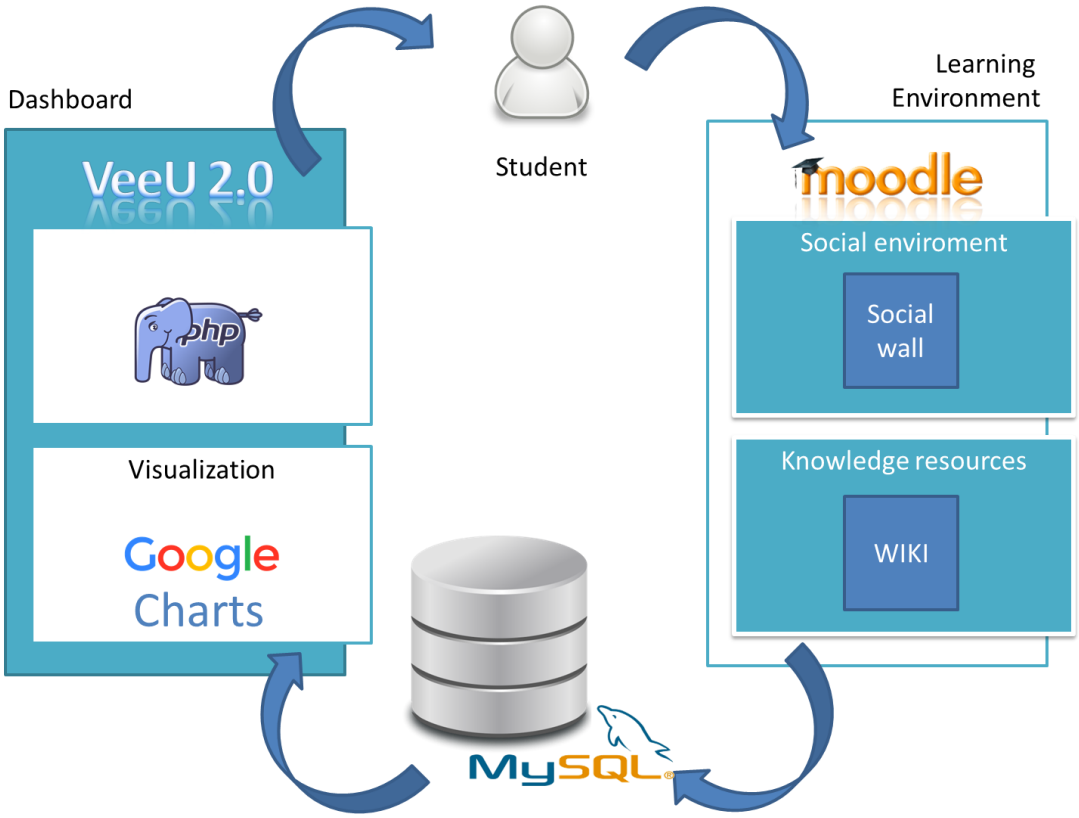


Figure 1. The Social Learning Environment

All the interactions occurring in the social learning environment are traced and stored in the Moodle database that is the data source of our dashboard. The trace-based approache, indeed, is one of the best method used in literature to detect users engagement, since they gather a lot of information without interfering with the learning activities [29] [30].

B. Engagement indicators

Engagement has been deeply studied in technology enhanced learning research from a wide range of perspectives: it is commonly recognized as a multidimensional and multifaceted construct but the definition of engagement is context dependent. In the specific context of social learning environments the engagement is strictly related to the activities performed in the community and to the users' participation in different kinds of activities.

First of all, it is necessary to classify the level of user's participation in the community, from a lower level, or peripheral participation [31], mainly consisting of reading resources and limited interactions with other users, to a higher level, or more active participation, consisting of activities such as writing comments on other users posts, sharing knowledge resources enabling a more responsible contribution to the knowledge of the community.

According to this perspective, in our research the engagement indicators were defined as follow:

- Passive interactions
  - Number of likes received by a post or a comment
  - Number of wiki pages read
- Active interactions
  - Number of posts published on the social wall
  - Number of comments shared on the social wall
  - Numbers of created wiki pages
  - Number of revised wiki pages

Furthermore, the next step is to identify the subject of interest of the user's participation. As defined in [32] the indicators of engagement are based on the participation both in social life of the community (interaction with other participants) and in the knowledge building activities (interaction with the knowledge resources).

According to the course structure we defined the engagement indicators as follow:

- Social Interactions
    - Number of posts published on the social wall
    - Number of comments shared on the social wall
    - Number of likes received on a post or a comment
- Knowledge Interactions
    - Number of wiki pages read
    - Number of created wiki pages
    - Number of revised wiki pages

All the indicators are calculated from the traces collected in the Moodle database: in particular standard Moodle log tables has been used for the indicators based on the Wiki while the SocialWall log tables have been used for the related indicators.

The basic indicators are collected at daily intervals and are then aggregated weekly or monthly according to the required level of analysis. The common aggregation level makes the indicators comparable over time and allows the timely evolution of engagement to be assessed. Then, as depicted in Fig. 1, the calculated indicators are presented through Google Chars using the most appropriate visualization as detailed in the next section.

C. Visualization

Our first goal was to visualize information about the entire class in order to understand the behavior and the level of participation of each student. Then we use our first pair of indicators to display the position of each student on a scatterplot. The students' position in the two-dimensional area is defined according to the number of Passive Interactions in a week on the x-axis and the number of Active Interactions in a week on the y-axis (Fig. 2). This visualization will provide a quick overview of the current status of the entire class, thanks to the scatterplot ability to display trends and relationship in a cloud of points [33]. Furthermore, outlier students will be easily identified but, in order to understand the reason of their poor participation, a detailed visualization is required.

To analyze the behavior of single student a visualization of engagement trends over time is needed. Thus, the second visualization (Fig. 3) depicts the details of a particular student in terms of trends and distribution. It is based on a linear visualization, a time series: two lines (one for each indicator) will display the trends of the engagement indicators over time for Passive and Active Interactions. In this case the indicators can be collected using daily, weekly or monthly intervals according to the required level of analysis.

Furthermore a pie chart will display the distribution of interaction in a selected time interval, between passive and active interactions, but also among the different components of each indicator (Number of likes received, Number of wiki pages read vs Number of posts published, Number of comments shared, Numbers of created wiki pages and Numbers of revised wiki pages).

After the analysis of the level of participation, the attention was focused on the kind of engagement. As a matter of fact, the more passive or active attitude of a student is not enough to evaluate the students' engagement.

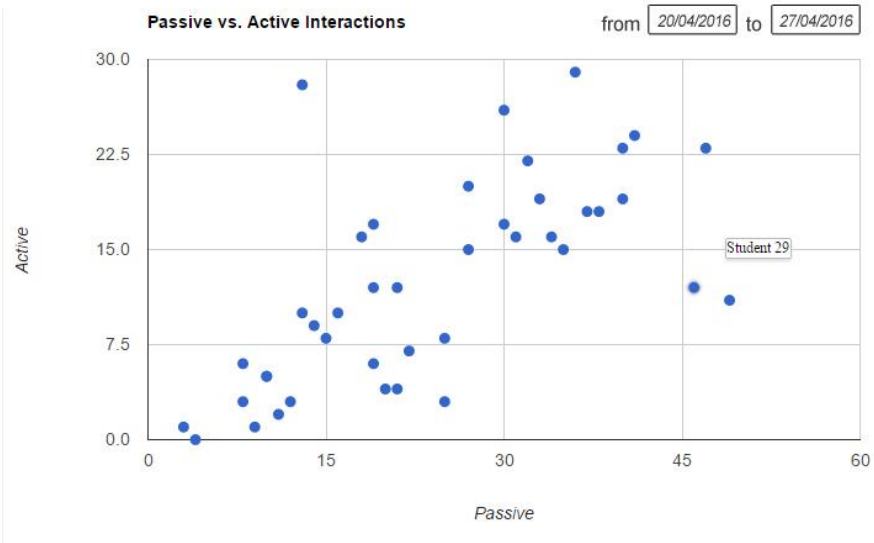


Figure 2. First visualization:
Scatterplot of Passive vs Active Interactions

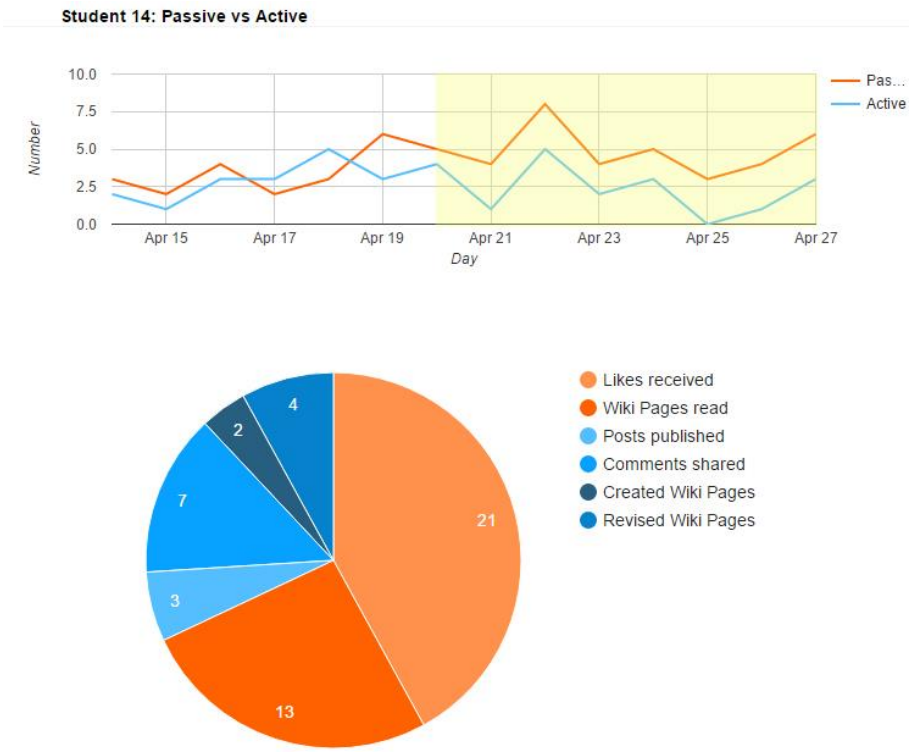


Figure 3. Second Visualization:
Trends and distribution of Passive vs Active Interactions for a single student

The third visualization (Fig. 4) is a scatterplot that, in this case, displays the students' position according to the number of Social Interactions in a week on the x-axis and Knowledge Interactions in a week on the y-axis. This kind of visualization will provide a quick overview about the subject of interests of student's engagement allowing the identification of purely socializer users or users actually involved in both Social and Knowledge Interactions.

A detailed visualization of a particular student will be useful in this case too. The fourth visualization (Fig. 5) presents the details of a particular student in term of trends and distribution with a time line and a pie chart built with this new pair of indicators.

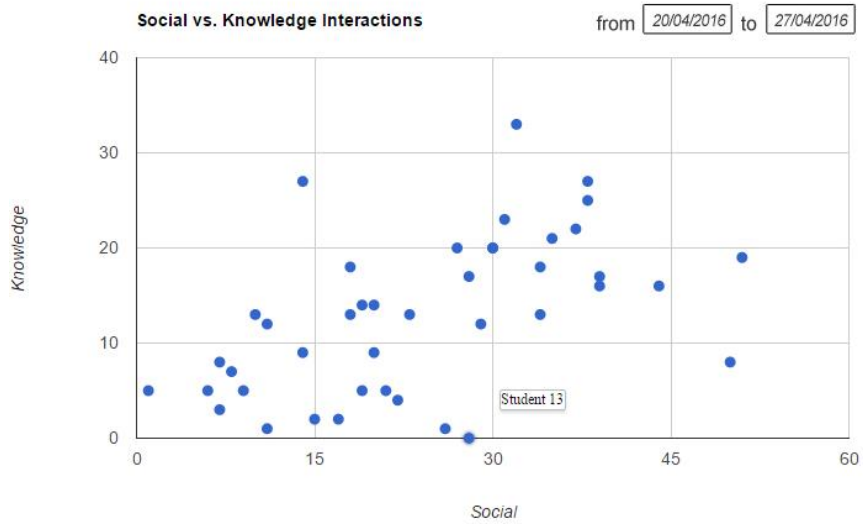


Figure 4. Third visualization:
Scatterplot of Social vs Knowledge Interactions

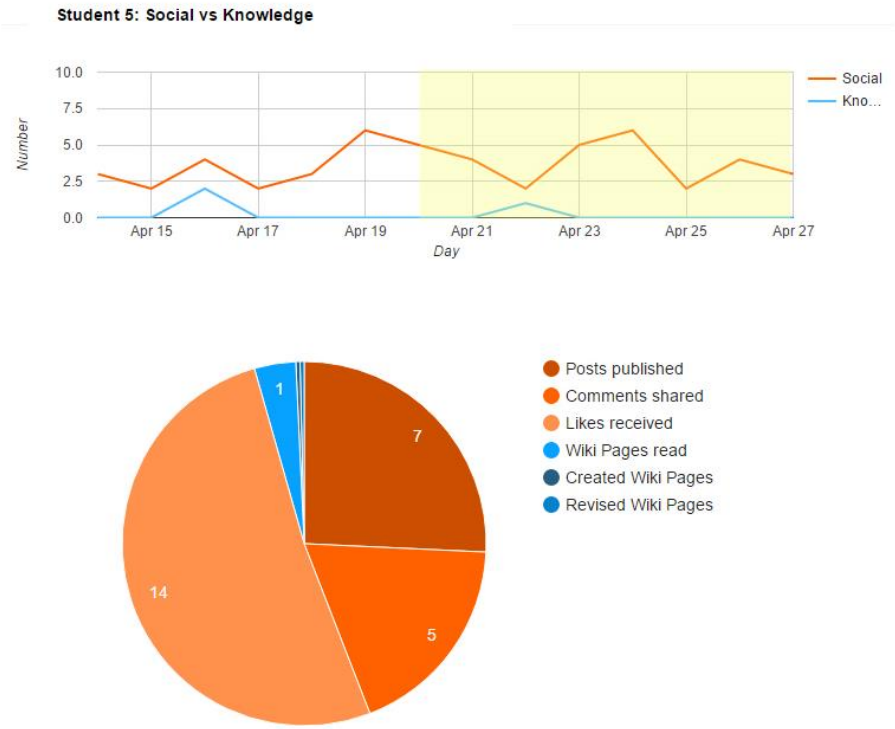


Figure 5. Fouth Visualization:
Trends and distribution of Social vs Knowledge Interactions for a single student

IV. CONCLUSION AND FUTURE WORKS

Our research aims to address the needs of information about the learning effectiveness in social learning environment, using engagement measures.

To this end, activity monitoring tools have become vital in distance education, thanks to the increasing availability of data collected by tracking the online activities. Information Visualization techniques, taking advantage of visual perception skills, are powerful tools to present a large amount of data by transforming traces in information. An accurate dashboard design is essential to provide students and teachers with a decision support tool that (1) promotes the awareness of what is happening in the learning process, (2) encourages personal reflection and achieving goals.

Data and reports usually provided by learning management systems are usually limited to measurement of activities and performances but there is a growing need to monitor less tangible aspects such as engagement in the social learning contexts. This is the main aim of the proposed dashboards, which is still under development. The approach adopted is an iterative process, which involves both researchers and social learning environment users. Researchers in the first design phase have involved teachers in order to define the most relevant indicators and to define the most appropriate visualization for the dashboards. The student were engaged in the pilot study, which involved two classes of undergraduate students from the Computer Lab courses held by the Department of Education, Psychology and Communication of the University of Bari Aldo Moro. Students are currently working in the social learning environment and the results of the beta test phase of the dashboard will be soon available. First results seem promising but further field test are required for both to measure the long-term effectiveness of dashboards in the learning process and to improve dashboard functionalities.

Future developments are under investigation. In particular, we are interested in monitoring the evolution of engagement indicators over time both for each single student and for the whole class. Static or dynamic time mapping approaches will be evaluated. Static visualization, such as small multiples, is more effective for analysis, whereas dynamic visualization, such as animation, was found to be more effective for presentation [34]. As stated by Chevalier [35], animations are commonly used (1) to replay history of events that occurred in a dynamic system allowing users to go back in time, and (2) to make activities and change visible at the same time, while keeping the user engaged thanks to the ability to attract user's attention, However, animation is not immune to fault such as higher cognitive load or perceptual effects of change blindness [36]. In order to overcome this drawbacks interactivity may be the key to overcome the difficulties of perception and comprehension [37] allowing users to explore time dimension, if needed, without distracting them from focusing on a single moment in time.

Furthermore, other indicators will be required to provide users with further insight of what is happening in the social learning environment. In particular, Social Network Analysis measures and social network visualization will allow us to

make explicit the relationships between users; Discourse and Content Analysis will allow us to explicit the topic of interest [38]. Even in this case interactive visualization will be useful to analyze and present the available data: for instance a tag-cloud visualization of topics will allow users to refine engagements indicator based on a most relevant or on a specific topic.

REFERENCES

[1] J. Kennedy, "Characteristics of Massive Open Online Courses (MOOCs): A Research Review, 2009-2012.," *J. Interact. Online Learn.*, vol. 13, no. 1, 2014.

[2] T. R. Liyanagunawardena, A. A. Adams, S. A. Williams, and T. Rekha Liyanagunawardena, "MOOCs: a systematic study of the published literature 2008- 2012," *Int. Rev. Res. Open Distance Learn.*, vol. 14, no. 3, pp. 202–227, 2013.

[3] W. Admiraal, B. Huisman, and O. Pilli, "Assessment in massive open online courses," *Electron. J. e-Learning*, vol. 13, no. 4, pp. 207–216, 2015.

[4] P. J. Muñoz-Merino, J. a. Ruipérez-Valiente, C. Alario-Hoyos, M. Pérez-Sanagustín, and C. Delgado Kloos, "Precise Effectiveness Strategy for analyzing the effectiveness of students with educational resources and activities in MOOCs," *Comput. Human Behav.*, vol. 47, pp. 108–118, 2015.

[5] E. Pesare, T. Roselli, V. Rossano, and P. Di Bitonto, "Digitally enhanced assessment in virtual learning environments," *J. Vis. Lang. Comput.*, vol. 31, pp. 252–259, 2015.

[6] G. Siemens, "Learning Analytics : Envisioning a Research Discipline and a Domain of Practice," *2nd Int. Conf. Learn. Anal. Knowl.*, no. May, pp. 4–8, 2012.

[7] D. Bajzek, W. Brown, M. Lovett, and G. Rule, "Inventing the digital dashboard for learning," *World Conf. Educ. Multimedia, Hypermedia Telecommun. 2007*, pp. 1084–1092, 2007.

[8] G. Hwang, "Definition , framework and research issues of smart learning environments - a context-aware ubiquitous learning perspective," *Smart Learn. Environ.*, vol. 1, pp. 1–14, 2014.

[9] K. Verbert, S. Govaerts, E. Duval, J. L. Santos, F. Van Assche, G. Parra, and J. Klerkx, "Learning dashboards: An overview and future research opportunities," *Pers. Ubiquitous Comput.*, vol. 18, no. 6, pp. 1499–1514, 2014.

[10] S. Govaerts, K. Verbert, E. Duval, and A. Pardo, "The student activity meter for awareness and self-reflection," *Chi*, pp. 869–884, 2012.

[11] S. Govaerts, K. Verbert, J. Klerkx, and E. Duval, "Visualizing activities for self-reflection and awareness," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 6483 LNCS, pp. 91–100, 2010.

[12] T. D. Loboda, J. Guerra, R. Hosseini, and P. Brusilovsky, "Mastery Grids : An Open Source Social Educational Progress Visualization," pp. 235–248, 2014.

[13] J. L. Santos, K. Verbert, S. Govaerts, and E. Duval, "Visualizing PLE usage," *CEUR Workshop Proc.*, vol. 773, pp. 34–38, 2011.

[14] J. L. Santos, S. Govaerts, K. Verbert, and E. Duval, "Goal-oriented visualizations of activity tracking : a case study with engineering students," *Learn. Anal. Knowl.*, pp. 143–152, 2012.

[15] J. L. Santos, K. Verbert, and E. Duval, "Empowering students to reflect on their activity with stepup!: Two case studies with engineering students.," *CEUR Workshop Proc.*, vol. 931, pp. 73–86, 2012.

[16] J. L. Santos, K. Verbert, S. Govaerts, and E. Duval, "Addressing learner issues with StepUp!," *Proc. Third Int. Conf. Learn. Anal. Knowl. - LAK '13*, p. 14, 2013.

[17] K. Verbert, E. Duval, J. Klerkx, S. Govaerts, and J. L. Santos, "Learning Analytics Dashboard Applications," *Am. Behav. Sci.*, no. February, p. 0002764213479363–, 2013.

[18] V. Rivera-Pelayo, V. Zacharias, L. Müller, and S. Braun, "Applying quantified self approaches to support reflective learning," *Proc. 2nd Int. Conf. Learn. Anal. Knowl. - LAK '12*, p. 111, 2012.

[19] G. Conole and K. Fill, "A Learning Design Toolkit to Create Pedagogically Effective Learning Activities," *J. Interact. Media Educ.*, vol. 8, pp. 1–16, 2005.

[20] M. May, S. George, and P. Prévôt, "TrAVis to Enhance Students' Self-monitoring in Online Learning Supported by Computer-Mediated Communication Tools," *Ijcisim*, vol. 3, no. May 2016, pp. 623–634, 2011.

[21] B. Yousuf and O. Conlan, "Enhancing Learner Engagement through Personalized Visual Narratives," *2014 IEEE 14th Int. Conf. Adv. Learn. Technol.*, pp. 89–93, 2014.

[22] B. Yousuf and O. Conlan, "VisEN: Motivating Learner Engagement Through Explorable Visual Narratives," in *Design for Teaching and Learning in a Networked World*, Springer, 2015, pp. 367–380.

[23] C. Coffrin, L. Corrin, P. de Barba, and G. Kennedy, "Visualizing patterns of student engagement and performance in MOOCs," *Proceedins Fourth Int. Conf. Learn. Anal. Knowl. - LAK '14*, pp. 83–92, 2014.

[24] R. F. Kizilcec, C. Piech, and E. Schneider, "Deconstructing Disengagement : Analyzing Learner Subpopulations in Massive Open Online Courses," *Lak '13*, p. 10, 2013.

[25] P. Di Bitonto, E. Pesare, V. Rossano, and T. Roselli, "Smart learning environments using social network , gamification and recommender system approaches in e- health contexts," in *Smart Education and Smart e-Learning*, Springer International Publishing, Ed. 2015, pp. 491–500.

[26] B. Schreurs, M. de Laat, C. Teplovs, and S. Voogd, "Social Learning Analytics applied in a MOOC-environment Authors," *eLearning Pap.*, vol. 36, no. January, pp. 45–48, 2014.

[27] R. Carrillo;, E. Lavouè, and Y. Prié, "Towards Qualitative Insights for Visualizing Student Engagement in Web-based Learning Environments Rubiela Carrillo The Dimensions of Student Engagement," *Proc. 25th Int. Conf. Companion World Wide Web*, pp. 893–898, 2016.

[28] S. Few, *Information Dashboard Design: Displaying data for at-a-glance monitoring*. Analytics Press, 2013.

[29] J. M. Kivikangas, G. Chanel, B. Cowley, I. Ekman, M. Salminen, S. Järvelä, and N. Ravaja, "A review of the use of psychophysiological methods in game research," *J. Gaming Virtual Worlds*, vol. 3, no. 3, pp. 181–199, 2011.

[30] P. Bouvier, K. Sehaba, and É. Lavoué, *A trace-based approach to identifying users' engagement and qualifying their engaged-behaviours in interactive systems: application to a social game*. 2014.

[31] J. Lave and E. Wenger, "Situated learning: Legitimate peripheral participation," *Learn. doing*, vol. 95, p. 138, 1991.

[32] E. Pesare;, T. Roselli, and V. Rossano, "Engagement in social learning: Detecting engagement in online communities of practice," in *International Conference on Applied Human Factors and Ergonomics*, 2016, p. in press.

[33] M. Friendly and D. Denis, "The eraly origins and development of the scatterplot," *J. Hist. Behav. Sci.*, vol. 41, no. 2, pp. 103–130, 2005.

[34] W. Aigner, S. Miksch, W. Muller, H. Schumann, and C. Tominski, "Visualizing time-oriented data-A systematic view," *Comput. Graph.*, vol. 31, no. 3, pp. 401–409, 2007.

[35] F. Chevalier, N. H. Riche, C. Plaisant, A. Chalbi, and C. Hurter, "Animations 25 Years Later : New Roles and Opportunities," pp. 280–287, 2016.

[36] Y. Albo, J. Lanir, P. Bak, and S. Rafaeli, "Static vs . Dynamic Time Mapping in Radial Composite Indicator Visualization," pp. 264–271, 2016.

[37] M. Tversky, B., Morrison, J. B., & Betrancourt, "Animation : can it facilitate ?," *Int. J. Hum. Comput. Stud.*, pp. 247–262, 2002.

[38] S. Buckingham Shum and R. Ferguson, "Social Learning Analytics," *Soc. Learn. Anal.*, vol. 15, no. 3, pp. 3–26, 2012.

Is e-learning ready for big data? And how big data would be useful to e-learning ?

Paolo Maresca[1], Andrea Molinari[2,3]
[1]Department of Ingegneria Elettrica e Tecnologie delle Informazioni (DIETI) – Università Federico II di Napoli (ITALY)
[2]School of Industrial Engineering and Management, Lappeenranta University of Technology (FINLAND)
[3] Department of Industrial Engineering - University of Trento (ITALY)

Abstract **The paper presents an overview of possible application fields of big data to the Technology-Enhanced Learning (TEL), with the many different facets this could imply. Many are the benefits for e-learning when approaching the collection of data, especially when e-learning is delivered in a life-long learning perspective. All these benefits could impact the future of eLearning, by revolutionizing the way we analyze and assess the eLearning experience. On one side, we present our experience in enriching the persistence layer of an LMS with a deeper log system on users' actions, in the perspective of collecting volumes of data compatible with big data tools and technologies, while highlighting some related issues. On the other hand we will deal with the first applications of cognitive systems that are responsible for catalysing the big data in analytics aimed at e-learning activities.**

Keywords: e-learning, big data, LMS architecture, cognitive systems

I. INTRODUCTION

Today, big data is one of the buzzwords that IT researchers and specialists use everywhere. So we have seen various "waves" of attention and "hype" on technologies and solutions related to Technology Enhanced Learning, but many of these hypes have revealed to be simply another buzzword and have not been consolidated. Today's buzzword that has been sweeping the world for a few years but has only just started appearing more commonly in eLearning is Big Data [10].

The term "Big data" refers to the huge amount of data coming from many different data sources, that become too large, complex and dynamic for any conventional data tools to capture, store, manage and analyse. Big data approaches and technologies interest many different application fields, and we will find "big data …." in front of any data analytics tool. Most of the problems were related with data warehouses, technologies aimed at supporting decisions based on reconciliation of the different databases of the organization.

The research activities in big data are aimed to find faster and more scalable solutions to store and process all data collected, instead of using traditional data warehousing approaches that are expensive, hard to design and to implement. Big Data therefore introduced two issues: how to address the problem of storing such a large amount of data, and how analytics tools could be created for the problem of analysing these huge datasets.

The paper will present the experience of the authors in designing and implementing a mechanism that will generate flushes of data from the persistence of the virtual learning platform created for managing tens of thousands users in our region, both from Academy and Industry. After the design of the solution, we have implemented a traditional internal method for managing the first data source for big data, i.e., dedicated logs of the platform. With a potential of more than 70.000 users, the volume at the moment are still in the range of the high level of database application, with approximately ten of thousand users using every day the platform for their e-learning tasks.

The tasks are recorded both for educational and security reasons, being the logging tasks involved in the recording of mouse clicking and in SCORM(Shareable Content Object Reference Model)-based material logging. Being SCORM logging not enough for some educational paths, according to the known limitations of SCORM standard, we decided to implement a meta-SCORM service, where more SCORM packages could be used in an educational path. In this scenario, we have logs of the platform for clicks and users' actions, logs from the SCORM player, and logs from the educational path services.

In general, there are several elements of data gathering and manipulation inside e-learning platforms [11] that could push this application field towards tools and techniques typical of Big Data:

- Traditional Weblogs, being the application a web-based software;
- internal logs of usage of the platform, the so-called "digital breadcrumbs", that track the learner's journey throughout the entire learning experience;
- Mobile logs, where data about mobile learning actions are collected;
- Service logs, users' actions on the different elements of the platform like documents, forums, blogs, FAQ etc.;
- logs from the SCORM player, normally an external entity respect to the core services of

DOI reference number: 10.18293/DMS2016-042

the platform, with the records of the SCORM objects' execution;

- Tin-CAN API calls, in case the platform is connected or acting as a Learning Record Store (LRS).

Recently, a further set of new ideas in e-learning could increase the need of a structural change of Learning Management System (LMS) architecture towards approaches and technologies connected with big data:

- Massive Open Online courses (MOOCs), by definition a generator of high volumes of data

- Life-long learning, an old buzzword of e-learning that is still valid and interesting and, most of all, is another generator of big data, specifically along time;

- Serious games that will use materials inside the platform, thus generating a relevant dataset related with users' performances;

- Big Data will change the way we approach e-Learning design by enabling developers to personalize courses to fit their learners' individual needs [14].

The paper will present an overview of problems and opportunities related with the introduction of a big data approach to e-learning, both in the software architecture of the platforms and in the approach of e-learning stakeholders to this discipline. The paper is divided as follows: section 2 will present the potential application fields of big data technologies and approaches in e-learning. Section 3 will illustrate one specific example of big data in e-learning, i.e., the change in the persistence layer of a software platform for e-learning, while section 4 will present the use of big data technologies, specifically IBM "ask Watson" solution to e-learning settings.

II. Big Data And E-Learning: The Potential Application Fields

There are a number of reasons why big data may, very well, revolutionize the eLearning industry. First and foremost, it will allow eLearning Professionals to customize the eLearning experience to provide learners with more effective, engaging, and informative eLearning courses and modules. Big data also has the potential to impact the future of eLearning by:

a) Offering invaluable feedback.

While online surveys and discussions may offer feedback regarding the effectiveness of eLearning courses and modules, big data gives to eLearning professionals the chance to receive invaluable feedback that can be used to pinpoint where the learner, and the eLearning course itself, may need to be improved. For instance, if a learner is able to look at an analysis of where he/she fell short while taking the eLearning course, he/she can then figure out how to correct the issue moving forward. At the same time, if the online facilitator observes that the majority of the learners struggle with a particular module or assignment, he/she can make proper adjustments to improve learners' performance.

b) Allowing eLearning professionals to design more personalized eLearning courses.

If eLearning professionals are given the opportunity to know how their learners are acquiring information and what works best for them, in terms of content and delivery, then this will result in more personalized and engaging eLearning courses. As such, modules can be custom tailored to meet the individual needs of the learner, which will offer a high quality and meaningful learning experience.

c) Targeting effective eLearning strategies and eLearning goals.

Big data in eLearning gives us an inside look at which eLearning strategies are working and which ones aren't necessarily helpful in terms of eLearning goal achievement. For example, you can determine which eLearning courses are contributing to skill development and which eLearning modules or elements may be irrelevant. As such, you can then devote resources to the aspects that are working, so that the learners can receive the preparation they need to fulfill their career goals.

d) Tracking learner patterns.

With big data, eLearning professionals gain the rare ability to track a learner throughout the entire process, from start to finish. In other words, you can see how well they performed on a test, or how quickly they finished a challenging eLearning module. This helps you to pinpoint patterns that will not only enable you to learn more about the learning behaviors of the individual learner, but your learners' group as a whole.

e) Expanding our understanding of the eLearning process.

As eLearning professionals, it's essential that we learn as much as possible about how learners acquire and digest knowledge. Big data gives us the chance to gain an in depth understanding of the eLearning process and how the learners are responding to the eLearning courses we are delivering to them. We can even figure out which time of day they learn most effectively or which delivery methods allow them to retain information most efficiently.

This information can then be used to take our eLearning strategies to the next level. To make a long story short, embracing Big Data reshapes eLearning and it's a win-win situation for both the eLearning professionals and the learners.

III. ONE TESTBED FOR BIG DATA GENERATION: THE "ONLINE COMMUNITIES" PLATFORM

The source of inspiration for big data analysis in e-learning is a virtual communities management platform entirely and autonomously created by our research team, starting since 1998. The approach used in most of the communities managed by the platform regards what is called "blended" approach, i.e. an e-learning mixed between frontal and online education training, asynchronous and synchronous, with online tutoring and frontal work sessions, all supported by our "Online Communities" system, of which, around the end of the 90s, the Faculty of Economics of the University of Trento has decided to adopt, followed by other public and private institutions.

Currently, "Online Communities" (OC from now on) is mainly used outside the university campus, serving approximately 50.000 users from different public and private customers against approximately 15.000 students in our University [16].

OC is a dynamic web application, based on the metaphor of the virtual learning community, which ensures the cooperative organization of work in groups of users called "Community". A virtual community is defined as a space of communication shared by a group of people, not only related to educational aspects. Every community has at least one coordinator and participants that are not anonymous. It is natural to imagine a virtual community as an aggregation of individuals made possible thanks to computers; an extension in the virtual learning environments is the class in which the courses take place.

The system is designed from the ground up within the Laboratory of Maieutic working group - Department of Industrial Engineering of the University of Trento, and is able to support the needs of a broad group of users (teacher, student, tutor, lecturer, external consultants, supervisor, dean, counsellor, secretary, board member etc.), customizable within the context in which the system is used (for example in a business organization we will have different roles respect to the university, as president, secretary, director, administrative, board of director etc.).

The participants in the system are not anonymous, and have a number of roles; each role brings with it specific rights and duties. Therefore, the enrolled users participate in a series of communities, fulfilling different roles in each of them. The communities are also characterized by a series of events that correspond to the active involvement of members in different moments. OC was released outside of test environments in 1998 as a working prototype, then reached through different evolution steps its maturity in early 2005, but counting on a long experimentation on a limited number of courses started at the end of 1998.

Since September 2005, the system is in operation at the University of Trento, involving tens of thousands of users, and has become an everyday tool for the teaching of many teachers. It represents the technology infrastructure used officially by some faculties of the university of Trento for testing and supporting new forms of teaching based on the techniques and the methods of e-learning. In 2007, the Autonomous Province of Trento decided to adopt it as its platform to deliver e-learning for its approx. 20.000 users. Then the Chamber of Commerce of Trento and the Chamber of Commerce of Bolzano adopted the platform and extended it to the affiliated enterprises, where we have just for the Province of Trento approx. 55.000 enterprises with the respective users.

The platform enables the members of any community to collaborate using multiple computer services (blog, wiki, chat, video conferencing, forums, file sharing, diaries , diaries, etc.). These services are used in areas collaboratively, both in the training, given the apparent closeness between the two environments. It 'also clear that, talking about continuing education and lifelong learning, the boundary between training and collaboration is increasingly being blurred when dealing with adults involved in the workplace. The logic of a virtual community platform considers any combination of people, brought together in a virtual place (hence "virtual community") for various purposes.

From our analysis, at the moment the interaction of users generate 2GB of data per month only for the users' actions log, just for 2.000 users/day. It's clear that the overall picture could become much more compelling for any software platform where the collection of data could stimulate and support DSS from the top management of the institution, or where hundreds of thousands participants follow intensive MOOCs.

Most of this data are coming from the first data source we used in our experimentation due to its affinity with big data data sources, the aforementioned "Actions". It collects all data coming from users' interactions with any OC object or service. It is a sort of "sensor" introduced inside the source code of the platform in any place is needed the software to capture an "action" from the user interface. This is of course a relevant enrichment of the logs recorded by the web application server, and has been used for many different purposes.

Due to volume issues, the system at the moment is blocked on collecting only some types of events, to a certain granularity defined by the system administrator. This choice has not been a design choice, but a performance-related one. From the early experimentations, it was clear that the amount of data could have compromised the capacity of the DBMS to stand data acquisition pace and volumes: a typical "Velocity" and "Volume" big data problem.

Tasks are recorded both for educational and security reasons, being the logging tasks involved in the recording of mouse clicks and SCORM-based material actions. Being SCORM logging not enough for some educational paths, according to the known limitations of SCORM standard, we decided to implement a meta-SCORM service, where more SCORM packages could be used in an educational path, mixing them with other services provided by the platform (like Wikis, FAQs, forums etc.) in a unique view called "educational path". In this scenario, we have logs of the platform for clicks and users' actions, logs from the SCORM player, and logs from the educational path service.

From the potential amount of data generated by these actions, we started to imagine our big data approach to the platform.

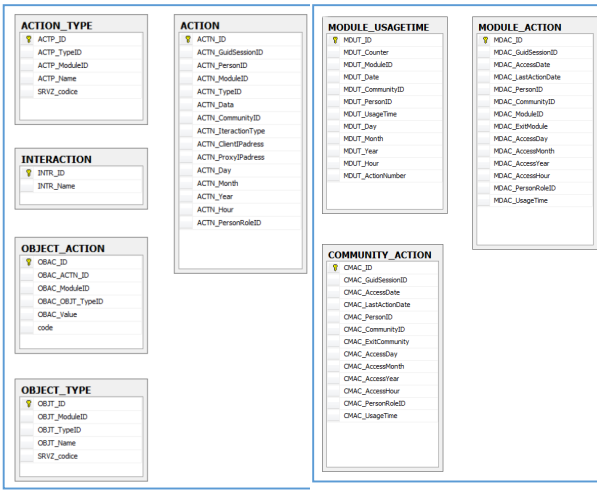


Figure 1. data structures for big data analysis in "Online Communities"

IV. Big Data As Support To Learning Management And Learning Process

One of the domains in which you can take on the challenge of big data and the cloud is the *management of learning processes* as well as the *customization of the learning management*. The *customization of learning processes* can be accomplished employing the cognitive system as the IBM Watson [2] according to a paradigm known as to ***ask Watson,*** with which you establish a triage *Watson-student-teacher* that will be explained in detail in the following paragraphs.

Thanks to the use of cognitive computing, classes become more intelligent, not only for the content they make available to the students but for the possibility of using student's data to figure out how to grow the class. The observation period may be as long as much as they are in the school and beyond[3].

These systems can also help teachers deal with the problems of the students such as school dropout, credits recovery etc., thus allowing a process that aim at learning *customization*. The paradigm is innovative and is based on the fact that the teacher asks the cognitive system to "talk" with the student in order to understand which are the strengths and weaknesses to be reported to the teacher who then will decide which course materials should be delivered to overcome the identified learning problem. This feature is known as ***Ask to Watson*** and represents the aforementioned ***triage*** *student-Watson-teacher* which represents the true individual learning customization.

As an example, you can ***ask Watson*** to discover the learning mode of the student (kinaesthetic, visual, auditory) in order to deliver the appropriate training material. Or you can ***ask Watson*** "to prepare a small subclass for the Olympics of mathematics or computer science" and then provide the material and proper training to achieve particular objectives challengers. But the most important thing is the interaction with Watson that occurs in spoken language.

To date, Watson interacts in English, Japanese, Spanish and Portuguese, it learns from the students using the features outlined in the previous paragraph. In addition, the teaching material is densely developed with HTML tags and hypertext, so that you do not need to do complex searches and the concepts are often proposed using games and flowing into the domain appropriate to young adults.

There are many existing projects in this area around the world. Georgia's Gwinnet County public schools, is the 14th most populous district of the USA and is one of the first college that has pioneered the use of Watson. Watson has the task of "*identifying similarities in how students learn and predicting performance and learning needs*", while "*specific content and teaching techniques can be aligned to each of the district's 170,000 students to Ensure the best learning experience*". In particular, the paradigm is quite complex and is constituted by recursive cycles that consist of the following 6 steps:

i) Watson identifies weaknesses and strengths of every students;
ii) Watson recommends behaviour and contents for students aligned to their skills and learning styles;
iii) Teacher selects appropriate learning path and Creates a progression plan for student;
iv) Students consume the recommended content from the material plan;
v) Teachers monitor the progress of the students and adjust the plan;
vi) Teachers use Watson to identify student skills attainment aligned to the standards defined.

More specifically, it is interesting to find out how they interact between Watson, the student and the teacher through the services it offers. The interaction occurs in spoken language and consists of three cognitive services: I) **Ask**, II) **discover**, III) **decide** delivered in sequence. In the first the student asks Watson to guide him, in the second Watson teaches the student how to derive the answers from the questions asked, in the third the teacher sees the results and delegates Watson for the adaptation of the training course. This mode of *assistance* is very promising and is bearing much fruit.

Other challenges have been launched in this field especially in the analysis of data and analytics [1][6][7]. For example, Wichita State University uses advanced analytics to predict potential students chances of success: 15% boost in registration. Hamilton County, Tennessee's Department of Education uses predictive analytics to improve student achievement causing the Following results: 8% increase in the graduation rate to 80% and 25% reduction in the annual drop rate. Seton Hall has used integrated marketing optimization solutions analysing social media to understand how students move through the recruitment process: 13% enrolled increase.

The novelty of this approach lies in the use of natural language (English, Spanish, Portuguese and Japanese) that allows Watson to discover the student's difficulties

and to report them to the teacher who chooses which educational materials to provide, and so on until you reach the training course objectives that have been established for that student. Training materials can be chosen on the basis of intellectual channels of young people and can be provided on the basis of the receptive attitudes of the students (auditory, visual or kinaesthetic).

At this stage, the interaction through natural language helps a lot especially if you are using two essential functions of cognitive system that are called speech to text and text to speech (see fig.2). The first transforms written text into speech, the second performs the inverse function. Based on these features, for example Sundararajan and Nitta [12], designed and realized a tutoring system for K12 students, intensively using interactivity, automatic generation of questions, and learning analytics. It is worth noting that these features, as all the others, are available through Platform-as-a-Service (PaaS) bluemix IBM and the access for Universities is a free of use under the IBM academic initiative agreement for six months.

Watson Ecosystem is one of the largest research organizations in the world, only in the IBM business intelligence industry spends 6 Billion of dollars for years since 5 years. Watson is accessible through a cloud open platform called Bluemix (see fig.2), has spread to 35 countries, with 18,000 programmers who work there and 500 people making their business all over the world. Watson currently has more than 30 services, 15 underlying platform technologies, and thanks to bluemix, enables the output of its API to anyone who requests it.

Figure 2. Watson and their services (boiler plate) inside Bluemix Paas

The big data analytics is the key of managing the learning processes [13][14]. This allows us to understand and observe the students careers during their training to provide them with assistance and suggestions about the composition of university curricula, management training curricular and extracurricular, insights and the advice to acquire specific skills. New generations of (special) applications that precisely interact with a cognitive system (and therefore are named ***cognos***) are starting to appear. An interesting experience of using cognitive computing in educational settings is the Watson Student Showcase [8] organized by IEEE jointly with IBM. Students were challenged in a competition where they had to develop cognitive apps, using the cognitive computing services included in Bluemix.

The cognitive computing and the internet of things (IOT) will be the real innovation of the future. This will have the disruptive implications in the way of teaching [9]. If you only think about how our laboratories will be managed and maintained, with everything on cloud and with students interacting with equipment and robots that will in turn interact with other things using simple interfaces and often through the voice. all this can be achieved by developing a new generation of cognitive app that uses the voice and therefore might call cognitive app (Cognos).

Cognos learn and communicate with other machines making it easy to manage the interaction between man and machine and from machine to machine. Consider the impact of all this on our teaching laboratories: electrical and electronic measurements, computer science, robotics etc. Labs may be automated, students will be able to use them at anytime from anywhere, without limitation of time. An approach of this type will revolutionize education, teaching and the contents. Of course cognitive computing as Watson and other can make this already possible [7].

Another interesting side is the bureaucratic management of teachers and students. For the management of students, the observation is not only limited to the period of University residence, but can be extended before (in high school) and after (in employment) in order to have a traceability of the careers of the skills and knowledge acquired. All this is achieved through the management of big data and through the management of platforms dedicated to this type of objective. Again cognitive systems [5] well compete through the analytics. Even the management of teachers is interesting through analytics, and more generally, many problems can be managed through analytics foreshadowing a smarter Universities [4].

The analytics can also be extended to networks of people who share a common goal. The ***nose of knowledge*** helped a lot in building a network of expertise and knowledge [7]. But this knowledge is often limited to students in a class network. This causes the potential of the group remain limited. If we were interested to boost the potential of this network group should be extended to accommodate the students' knowledge and skills from other classes, other universities geographically located distant from us. Precisely in this regard you may develop a ***nose of knowledge cognos app*** that extends the capabilities of another app developed to the Ohio state university.

In fact The **Ohio State University** created a web application for the Watson University Competition called the "**YouDoU**" [15] project that helps Ohio State students find activities based on their personal preferences using Watson technology. An extension of personal preferences are also represented by the skills

and knowledge acquired during their university studies that everyone wants to pool in a workgroup. This extension can be managed by ***nose of knowledge cognos app.*** Many analytics can be used. Analytics can be used to measure the reputation that each student acquires in each group, by virtue of the activity that takes place during the collaboration. This way we can dynamically reconfigure the network by combining the elements that achieve the highest score, excluding those that bear a lesser reputation of a certain threshold, etc.

Finally analytics can also be applied on Italian University's admission tests to identify the University course credits that each student must mature and then follow him in his career and give him a concrete and effective customized training support.

V. CONCLUSIONS

In this paper we presented some application scenarios of big data on e-learning. More generally, the authors think that the ability to think, learn and understand the world around them will be essential to compete in the near future both for universities and companies . All this will be possible by the use of big data. We are surrounded by big data, and it is natural to use them to do business through analytics. In Universities, it is natural to use big data to manage learning processes in an intelligent way and to help individual students to improve their skills.

From the perspective of companies, we are facing a very strong disruptive technology that will transform everything, i.e., cognitive computing. Technologies related with analytics, cloud , mobility, security are very important, but one question arise to which we have to find an answer: when will all companies be digitized enough to compete in the market and win the competition?

The point is that the process of digitization in a company is not among its primary goal, but constitutes the premises for further digital revolution: the cognitive IOT. The two drivers of this new technology will be invisible and cognitive data. Eighty percent of data is invisible and is not understandable, "obscure" so to say, because it has not a structure but is also stored somewhere. This includes video, music, news articles, research reports, social media posts, enterprise data system etc. The second driver is even stronger and is represented by cognitive that no longer requires programming.

Cognitive systems in fact require the use of spoken language and there are currently many existing applications, although at an experimental level, on sails, supply chain, research, child (with *dyno* cognos), bluematch (putting together positions, persons and curricula of each one in order to reach specific project objectives), education , cooking recipes, medicine, meteorology, etc.). The difference between the cognitive and the current systems is that systems understand, reason and learn. It 'a revolution in technology, business and teaching.

The revolution in teaching will be even stronger because it will change the paradigm, the educational materials and the laboratories, and this in turn will imply the change in the role of the teacher, that will have a co-coach role assisted by cognitive systems. The support to this new paradigm of teaching and learning will cause a circular mechanism, where more and more data will be needed. So the sooner the e-learning world will realize and implement big data inside their technologies and processes, the better for the whole e-learning movement.

REFERENCES

[1] E. AbuKhousa and Y. Atif (2014), "Big learning data analytics support for engineering career readiness," Interactive Collaborative Learning (ICL), 2014 International Conference on, Dubai, 2014, pp. 663-668.

[2] Banavar, G.S. (2015), *Watson and the era of cognitive computing,* in Proceedings of the 2015 IEEE International Conference on Pervasive Computing and Communications (PerCom), 95-95.

[3] Byrd, G. (2016), IEEE/IBM Watson student showcase, IEEE Computer, 49 (1), 102-104.

[4] Coccoli M., Guercio A., Maresca P. and Stanganelli L. (2014) "Smarter Universities: a Vision for the Fast Changing Digital Era", in Proceedings of 20th International Conference on Distributed Multimedia Systems (DMS2014)

[5] Coccoli M., Maresca P., Stanganelli L. (2012) "Cognitive computing in Education ", Journal of e-Learning and Knowledge Society (Je-LKS),Vol. 12, n. 2, May 2016, pp. 55 – 69.

[6] Ferguson, R. and Buckingham Shum, S. (2012). Social Learning Analytics: Five Approaches. LAK 2012. 23--33. DOI: 10.1145/2330601.2330616

[7] Gašević, D., Dawson, S., & Siemens, G. (2015). Let's Not Forget: Learning Analytics are about Learning. Techtrends, 59(1), 64--71. DOI: 10.1007/s11528-014-0822-x

[8] Holtel, S. (2014), *More the end of information over ow: how IBM Watson turn upside down our view on information appliances*, in Proceedings of the 2014 IEEE World Forum on Internet of Things (WF-I.T), 187-188.

[9] Coccoli M., Maresca P., Stanganelli L., and Guercio A. (2015) "Teaching Computer Programming in a Platform as a Service Environment", in Proceedings of 21th International Conference on Distributed Multimedia Systems (DMS2015)

[10] Riffai M. M. M. A., Duncan P., Edgar D. and Al-Bulushi A. H., "The potential for big data to enhance the higher education sector in Oman," 2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC), Muscat, Oman, 2016, pp.1-6. doi: 10.1109/ICBDSC.2016.7460346

[11] Siemens, G. & Long P. (2011) Penetrating the Fog: Analytics in Learning and Education. EDUCAUSE Review 46 (5), 2011, pp. 30-38.

[12] Sundararajan, S.C. and Nitta, S. (2015), Designing engaging intelligent tutoring systems in an age of cognitive computing, IBM Journal of Research and Development, 59(6), pp. 1-9.

[13] Vozniuk, A., Govaerts, S., & Gillet, D. 2013. Towards Portable Learning Analytics Dashboards. ICALT 2013, IEEE, 412--416. DOI: 10.1109/icalt.2013.126

[14] West, D. 2012. Big Data for Education: Data Mining, Data Analytics, and Web Dashboards. Governance Studies at Brookings, pp. 1-10.

[15] Youdou, https://www.youtube.com/watch?v=LiOI07dQVTI

[16] Colazzo, L.; Molinari, A.; Villa, N.; , "Collaboration vs. Participation: The Role of Virtual Communities in a Web 2.0 World," Education Technology and Computer, 2009. ICETC '09. International Conference on , vol., no., pp.321-325, 17-20 April 2009

Supporting Mobile Development Project-Based learning by Software Project and Product Measures

Rita Francese[1], Carmine Gravino[1], Michele Risi[1], Giuseppe Scanniello[2] and Genoveffa Tortora[1]

***Abstract*— Project-based learning is a kind of learning activity which has great educative effect, but which presents also several issues. In particular, if we consider an university course that requires the design and the implementation of a software project, may be difficult to estimate the number of hours that a team of students has to take to accomplish that project. There is the risk to underestimate the project (too difficult) or to overestimate it (too easy) with respect to the other projects of the same course and the amount of foreseen work hours. In this paper, we present the experience we gained in the adoption of Software Project and Product Measures for addressing the project size of projects performed during a Mobile Application Development course for Computer Science students at the University of Salerno. The course foresaw a project work conducted by students organized in teams. The goal of the project work was to design and develop an Android-based application with back-end for smart devices. Software estimation project measures are applied to some metrics extracted in the requirement analysis phase to get an estimation of the effort in terms of man/hours and consequently to adjust the project size by adding/reducing requirements. The metrics extracted from the projects of academic year 2013/14 have been used in the successive year for estimating the project effort and intervene on the project size variables.**

I. Introduction

Software development projects may often be late and overrun time and cost [1]. These can have a dramatic impact in case of Project-Based-Learning (PBL) activities, where it may occur that a student project is late or requires much more work with respect to the estimation. Effort understimation may produce the failure of the project and the team members risk to not pass the exam, while effort overstimation may produce a project not corresponding to the teacher expectations. The final evaluation can be excessive if the teacher does not correctly evaluate the project size, or students can have poor results, while they could have produced more with appropriate requirements.

Software estimation problems have been largely investigated in the literature. In particular, in [2] the problem of predicting software project and product measures in the domain of Android development has been addressed considering as sample mobile applications produced by students during a mobile application development course.

In this paper, we try to reduce the gap between the amount of hours required to develop a project and the real project effort in the domain of PBL for Android app development. We decided to try to estimate in the earlier phases of the projects the required effort and to resize the project requirements according to this suggestion. In particular, we exploited the results we got in [2], where we demonstrated that requirement analysis measures can effectively employed to estimate software project and product measures of a mobile app and estimations can be done early in the software development process. The metrics extracted from the projects of academic year 2013/14 have been used in the successive year to intervene on the requirements to adjust the project size.

The paper is structured as follows: in Section II we discuss the background concerning PBL and Software Estimation; Section III describes the learning experience related to the considered Mobile Application Development course; Section IV summarizes the effort estimation model we adopted and Section V describes how it is used to estimate student efforts and the results from its application. Finally, Section VI concludes the paper.

II. Background

Project-Based Learning (PBL) is a learning approach that structures learning around projects [3]. It involves both constructivisme [4], a learning theory in which learners actively construct their knowledge, and team-based learning [5], in which the learning activity is conducted by a team of people which collaborate to pursuit a common objective.

In particular, Vigotsky [6] puts in evidence the relevance of the social context and the interaction among pairs to learn when problems have to be solved. This overcomes the idea of learning activity where students passively absorb knowledge from their instructor. When student are motivated to cooperate they can be successful team member in real industry environments [7]. However, team work is not only work together, it requires the team members to plan their projects activities, to monitor the project progress, and to disciplinarly manage their work [8].

PBL let students cooperate to solve together problems typical of the job world. The involvement of students is in general higher than traditional classroom activities.

The adoption of PBL in Computer Science courses is growing [9], [10], [11], [12], [13] because it enables the students to acquire soft skills such as to be able to work in group as a team, to respect deadline, to take ones own responsibility, to be able to communicate. All these skills are considered very relevant for organizations. PBL promotes principles, methods and procedures similar to the ones adopted for developing real software products [14]. During the various project phases documentation is produced, following specific standards. This implies that students do not

[1]University of Salerno, Italy {francese, gravino, mrisi, tortora}@unisa.it
[2]University of Basilicata {giuseppe.scanniello}@unibas.it

DOI reference number: 10.18293/DMS2016-045

have to perform only coding activity, but also to acquire analysis and writing skills [10].

In [15] we described the learning experience related to the Mobile Application Development course, summarized in Section III. During this experience we collected several product and process metrics useful to assess the project and process quality. From this experience we gained the conviction that there was the need of following a project management approach which takes into account effort estimation for assigning correctly dimensioned projects to students' teams.

In the literature, estimation approaches have been classified in different way, see for example [16]. Generally, they belong to one of this types: *(i)* expert estimation, as software experts provides an estimate on the base of his experience; *(ii)* formal estimation model, a mathematical model is created on the base of historical data to quantify the effort; and *(iii)* a combination of the other two approaches. The estimation produces software measures which can be exploited for project management purpose.

Research studies put in evidence that the best approach does not exist, but the goodness of the results often depend on the context and the application domain [17], [18]. Also the development technology may influence the estimation accuracy [19].

Software estimation effort has been adopted in [20] for adjusting the size of database-oriented student projects, which followed a traditional waterfall development model. The prediction approach exploited a modified function point counting tables and use cases.

In this paper we apply the effort estimation model for mobile application development presented in [2], which is based on the metrics extracted from the projects of academic year 2013/14. This model has been conceived using the dataset of the experience presented in [15] and is detailed and discussed in Section IV.

III. The Learning experience

The Mobile Application Development (MAD) course is given at the University of Salerno since the academic year 2011/12. The course focus on the design and development of Android mobile applications. The course lasts 48 hours, divided in 24h lectures and 24h laboratory. It is estimated that for each hour of lab students have to perform at home two hours of study to implement their project.

In this paper we refer to two successive edition of the course: the first occurring in the spring of 2014 (described in [15]), and the second in the spring of 2015. The data got from the first edition have been used to develop an effort prediction model for the development of mobile applications, while in the second one we used the model to address the work of the student trying to get a balanced project work.

When the course begins, the project work is accurately motivated by each team by providing a *Project Proposal* document. The teacher evaluates this document, in which the objective of the project, the analysis of existing similar application, and a preliminary description of functional and non-functional requirements were provided. In particular, the teacher imposed the following non-functional requirements: the operating system is Android, the app had to interact with a remote server, communicating data through JSON; native device functionalities have to be exploited, including maps, GPS, sensors, phone call and SMS. Also the device rotation has to be managed and the app has to use SQLite to store data on the device. Games were admitted if they exposed back-end functionalities, such as account management, multiuser support, bonus management, and the upgrade of the app.

The team composition was freely chosen by students. We did not decide to assign randomly team members because the students had previous experience of project work in several courses and, at the last term, they know which were the classmates more appropriate for them. The students were asked to use GitHub[1] for the management activities. The lecturer created a GitHub account for each team of students. The templates of documents to be released at the end of the various development activities were made available in the GitHub repository of each team.

We established a schedule for each team of students. They were informed about deadlines and deliverables. The first deliverable was the project proposal, while the second the Requirement Analysis Document (RAD). Successively, the participants had to deliver the different releases of their mobile app and the final version of this app. To write the RAD, participants were asked to follow the template by Bruegge and Dutoit [21]. A RAD is used to document requirements elicitation and analysis. As for requirements elicitation, the software engineer specifies functional and non-functional requirements. Functional requirements are expressed as free-form text (a summary of functionality to be implemented) and then in terms of use cases narratives and use case diagrams. As for requirements analysis, the software engineer specifies object and dynamic models of the software. Each identified object of the problem domain is textually described, while relationships among objects are illustrated with class diagrams. Attributes and operations of problem domain objects are detailed only if needed. The behavior of these objects is documented in terms of sequence diagrams and state machine diagrams. These kinds of diagrams are exploited to specify complex behaviors of either use case or single objects, respectively.

We asked the students to follow an incremental prototyping development approach. The students were also asked to show app prototypes to the MAD lecturer before the conclusion of their project. At the end of the course a competition named *App Challenge* was conducted to whom participated members of International IT industries. This produced a good competition spirit among participants. The IT managers judged the final products of good quality and very near to real apps available on the market.

More detail on the teaching experience of the first edition are described in [15].

[1] https://github.com

TABLE I
VARIABLES DENOTING INFORMATION FROM RADS

Measure	Description
FR	Number of functional requirements
Act	Number of use case actors
UC	Number of use cases
Cla	Number of classes
SD	Number of sequence diagrams

TABLE II
VARIABLES DENOTING INFORMATION FROM SOURCE CODE OBTAINED BY THE UNDERSTAND TOOL

Measure	Description
McB	McCabe Cyclomatic complexity
Classes	Number of classes
Files	Number of files
Methods	Number of methods, including inherited ones
NL	Number of all lines
LOC	Number of lines containing source code
CLOC	Number of lines containing comment
STM	Number of statements
DIT	Depth of inheritance

IV. THE ESTIMATION MODEL

The estimation model has been created by analyzing the projects of the academic year 2013/14. In particular, we have considered the projects performed by students during the MAD course for Computer Science at the University of Salerno. Participants were originally 57 and were arranged in 27 teams. Data for 4 teams have been discarded due to incompleteness issues (e.g., lack of class diagrams or functional requirements) in the RADs the members of these teams produced. In the following of this section we describe how the model was built.

A. Variables

The dependent variable of interest for this discussion is the *Effort*, and can be computed as *the total effort to develop a mobile app expressed in terms of person/hour*.

In [2] initially were proposed two sets of independent variables. The first set, reported and described in Table I, collects variable obtained from the RAD or from the requirements, while the second set in Table II, collects variables from the source code gathered by exploiting the Understand[2] tool. These two sets of variables are useful to compare prediction accuracy of software measures obtained from RADs (RAD or requirements measures, from here on) against the accuracy of predictions obtained with measures gathered from source code (simply SC measures, from here on).

To explain how metrics have been collected, we provide an example of a project among those considered in our data analysis, *Archeotour*. Students of Archeotour team developed an Android application that provides information on archaeological sites and suggests tours considering the user's position and interests. Weather information is shown on demand.

The number of functional requirements (FR) of Archeotour is 8. It is obtained by counting requirements listed in the functional requirement section of the RAD. As an example, *"Select the site on the map and show its description, history and pictures"* is a functional requirement. A functional requirement can correspond to more than one use cases (UC). For example, the mentioned functional requirement is associated to the use cases *Show Site* and *Show Map* in Figure 1 showing one of the use case diagrams of the project. In this diagram, the number of use cases is 6, while the number of actors is 2. The number of actors (Act) is computed by counting the different actors that appear in all use case diagrams in the RAD. Similarly, the number of classes (Cla) is computed by counting classes in the class diagram of the same RAD. For example, the class diagram of Archeotour shown in Figure 2 is composed of 24 classes. The number of sequence diagrams (SD) is obtained by counting how many of this kind of diagrams have been specified in the RAD.

Some descriptive statistics (i.e., minimum, maximum, mean, median, and standard deviation values) of the independent variables are shown in Table III. For the dependent variables, descriptive statistics are also reported. We have also graphically shown the values for our dependent variables by the boxplots in Figure 3.

B. Estimation technique

The estimation technique adopted in [2] is StepWise Linear Regression (SWLR) technique [22], which explores the relationship between a dependent variable and one or more independent variables, providing a prediction model described by a linear equation:

$$y = b_1x_1 + b_2x_2 + ... + b_nx_n + c$$

where y is the dependent variable, $x_1, x_2, ..., x_n$ are the independent variables, b_i is the coefficient that represents

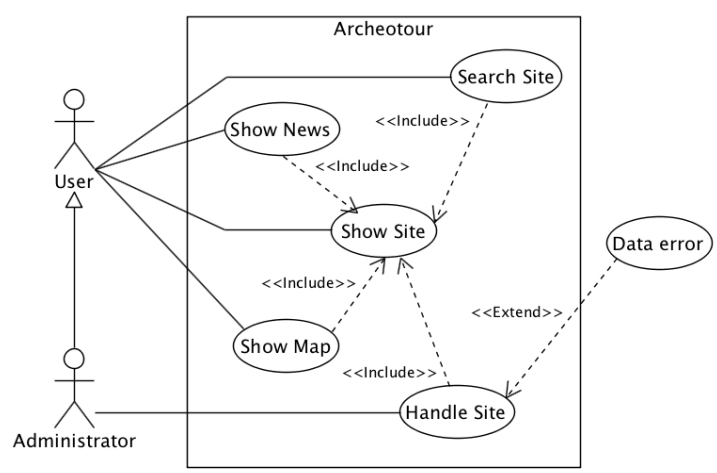


Fig. 1. A use case diagram of the Archeotur project.

[2]https://scitools.com

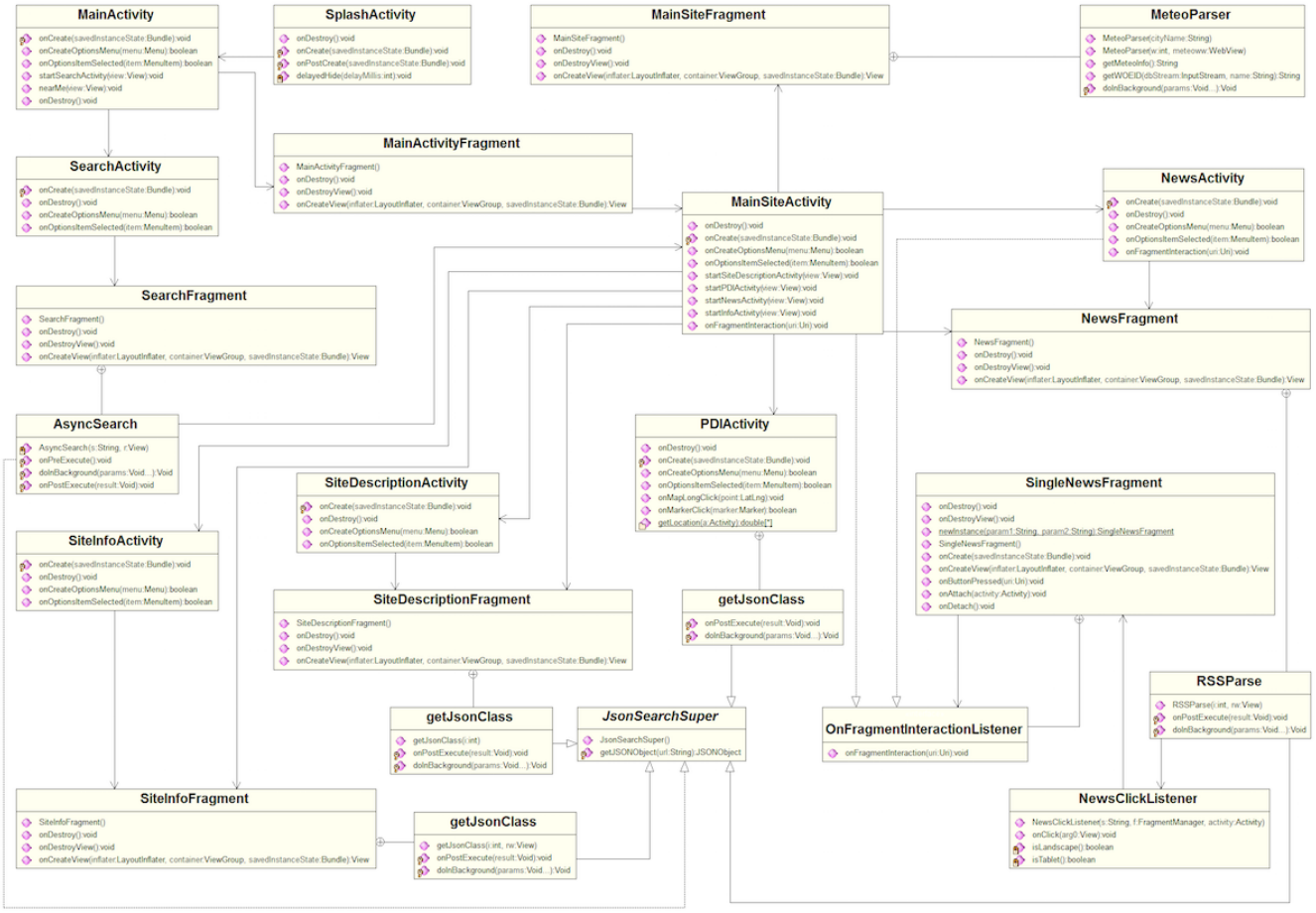


Fig. 2. The class diagram of the Archeotur project.

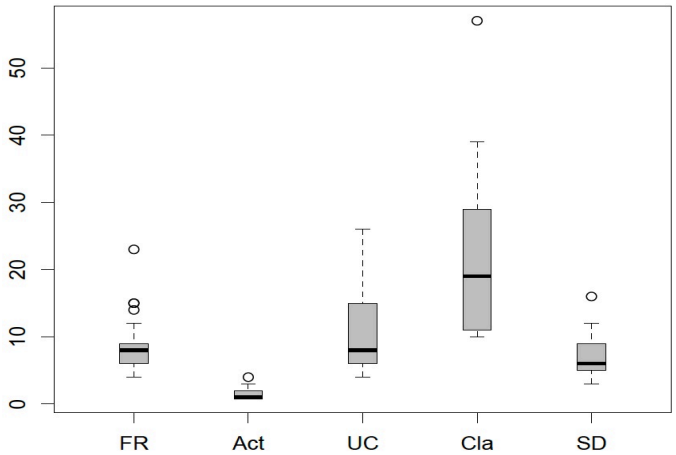
TABLE III

DESCRIPTIVE STATISTICS OF THE DEPENDENT AND INDEPENDENT VARIABLES CONSIDERED FOR THE ANALYSIS

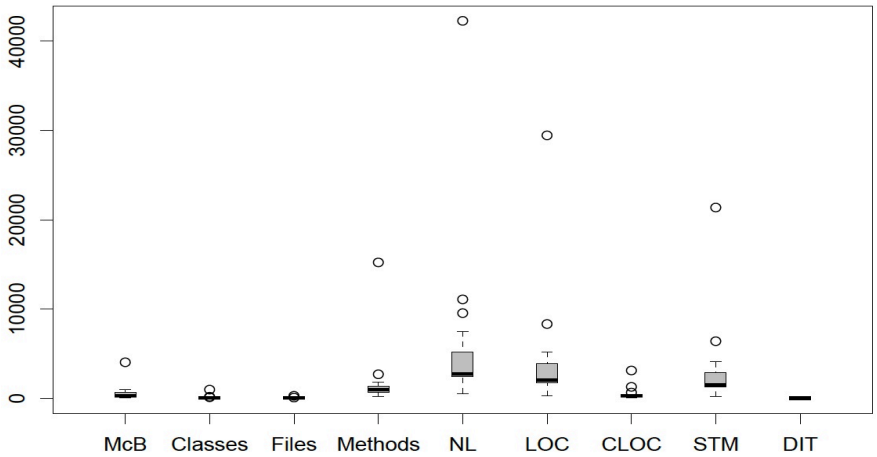
Variable	Min	Max	Mean	Median	Standard Deviation
FR	4	23	8.48	8	4.29
Act	1	4	1.59	1	0.8
UC	4	26	10.78	8	5.8
Cla	10	57	21.78	19	12.1
SD	3	16	7.07	6	3.025
McB	48	4030	517.52	282	747.91
Classes	12	967	89.22	54	178.62
Files	5	273	34.19	23	50.16
Methods	192	15222	1510.07	943	2795.71
NL	534	42287	5134.56	2740	7854.72
LOC	258	29456	3599.93	2037	5455.17
CLOC	12	3108	393.56	258	591.7
STM	163	21369	2714.44	1464	3969.11
DIT	2	4	2.44	2	0.58
Effort	30	113	58.82	55	21.04

the amount variable y changes when variables x_i changes 1 unit, and c is the intercept.

SWLR allows computing an equation in stages in which the choice of the independent variables is carried out by an automatic procedure. These variables can be chosen applying three approaches: forward, backward, or a combination of both [23]. The *forward* approach starts with no variables in the model. It tries out the variables one by one and includes them in the model if they are statistically significantly correlated with the dependent variable. The *backward* approach starts with all the variables and test them one by one. We remove the variables that are not statistically significant correlated with the dependent variable. We used here a combination of forward and backward approaches.



(a) RAD measures.



(b) SC measures.

Fig. 3. The boxplots of (a) RAD measure values, (b) and SC measure values.

At each step, this combined approach includes or removes variables one by one if they are or not statistically significant correlated with the dependent variable. We employed SWLR because this technique allows computing linear regression in stages and because it is widely used in the context of software prediction with appreciable results [17], [24], [25], [26].

To evaluate the goodness of fit of a model, we exploited the square of the linear correlation coefficient (i.e., R^2). This coefficient shows the amount of variance of the dependent variable explained by the model related to an independent variable. A good model should be characterized by a high R^2 value. We also considered the F value indicators and the corresponding p-value (denoted by Sign. F), whose high and low values, respectively, denote a high degree of confidence for the prediction.

C. The obtained effort prediction model

Before applying SWLR, we verified the following assumptions: *(linearity)* the existence of a linear relationship between independent and dependent variables; *(homoscedasticity)* the constant variance of error terms for all the values of independent variables; and *(normality)* the normal distribution of the error terms. We performed a log transformation of the input variables because the RAD measures were not normally distributed according to the results of a Shapiro test [27]. Furthermore, we performed the analysis of outliers, exploiting the Cook's distance and performed a stability analysis to eliminate influential observations [28].

The results of performed SWLR are summarized in Table IV. We can observe that the model built by RAD measures are characterized by a Sig. F value less than 0.05, thus the resulted model is significant. However, the obtained R^2 and F values are not so high.

In particular, the results revealed that best effort predictors include Cla, the number of classes in the RAD, and Act, the

TABLE IV
RESULTS OF SWLR FOR EACH DEPENDENT VARIABLE USING THE RAD MEASURES

Dependent variable	Independent variables	R^2	F	Sign. F (p-value)
Effort	Act	0.233	3.65	0.041
	Cla			
	Intercept			

TABLE V
EFFORT PREDICTED

ID	Number of participants	Act	Cla	Effort	Effort for participant
P1	4	2	31	56.53	14.13
P2	4	2	9	43.01	10.75
P3	3	1	7	49.13	16.38
P4	4	1	45	74.13	18.53
P5	4	2	22	52.40	13.10
P6	3	3	23	47.39	15.80

TABLE VI
REAL MEASURES AND EFFORT RESULTS

ID	Number of participants	Act	Cla	Effort	Effort for participant	Real effort for participant
P1	4	2	31	56.53	14.13	11.25
P2	4	2	35	58.06	14.52	24.75
P3	3	1	7	49.13	16.38	13.67
P4	4	1	30	67.77	16.94	17.25
P5	4	2	44	61.080	15.27	15.00
P6	3	3	23	47.39	15.80	16.33

number of use case actors, enabling the instantiation of the equation as follows:

$$Ln(\text{Effort}) = -0.272 * Ln(Act) + 0.221 * Ln(Cla) + 3.465$$

The final estimation model, when transformed back to the raw data scale, gives the following equation:

$$\text{Effort} = Act^{-0.272} * Cla^{0.221} * 31.96$$

A plausible justification for this outcome is that the number of classes in a RAD represents the basis for the next phases of the development process. That is, a developer uses these classes as starting point for implementation. Therefore, it seems reasonable that Cla provides useful information for an accurate prediction of the effort to implement apps. Furthermore, the number of use case actors give an indication of those interacting with the app, which is a crucial aspect for this kind of software.

V. MODEL APPLICATION

The successive edition of the MAD course was conducted with the same approach. Participants were 22. They were grouped in six teams. The data from the RAD have been collected as described in Section IV-A. We examined them and on the basis of the effort prediction measures, we proposed some little adjustments to the requirements concerning functionalities to be implemented as new classes. The number of participants for each team is reported in Table V, together with the number of actors and classes taken from the RAD and the estimated effort required for the project and for each participant. Let us assume that students work consecutively for three hours. This assumption is due to the fact that students follow also other courses and have only the afternoon free. Thus, the number of days to dedicate to the project implementation is 16. From these results it is possible to see that P1, P3 and P6 effort predictions are near to 16 days and we decided to leave them as they are; P2 and P5 have to be augmented to reach the required effort, while P4 has to be reduced.

The new versions of the RAD of these 4 projects have been analyzed and the results are in Table VI, together with the real effort. As it is possible to note, projects P4 and P5 reached a real effort near 16, while P2 dramatically exceed the estimation. Except P2, the other 5 projects respected the established time constraints.

VI. CONCLUSION AND DISCUSSION

In this paper we proposed to apply software estimation methods to the assignment of project works in mobile app development learning activities, specifically for Android application using a back-end server. The model has been created considering the data of the projects of year 2015 and has been applied to the RAD measures of the projects of year 2016. On the base of these measures, some adjustment of the project requirements have been done. The final estimation results on six projects were good for five of them, except for project P2, which doubled the required time to accomplish it. To better understand the reason why P2 was out of time we examined the type of the various applications. While P2 was a game with synchronization features, the others were dynamic web applications with Android interface and access to native features, e.g., an app which provides location-based search on swimming pools and events related to sea competitions. The servers-side of P2 was more complicated because of the need of connecting two players together in a single game session and transfer data messages between connected players. Thus, the metric values reported in this paper seems not to be effective for application with synchronization needs, considering the background of the students.

This approach may be useful to a teacher when it conducts the same type of course over the years. May be that when the project dataset grows collecting the data from different years the prevision may become more accurate. These considerations have to be confirmed by further work performed by

collecting data of the successive MAD courses. In addition, a much deeper analysis and setting of groups' composition should also be part of a larger, extended and deeper experiment which may consider also different variables, such as the composition of programmers' teams, gender, previous background and motivation. These could be other factors affecting final results which can be examined in future investigations.

ACKNOWLEDGMENT

We thank all the students to the two courses and the IT managers, who judged the apps these students developed.

REFERENCES

[1] K. Molokken and M. Jorgensen, "A Review of Software Surveys on Software Effort Estimation," in *Proceedings of the International Symposium on Empirical Software Engineering (ISESE)*, pp. 223–230, 2003.

[2] R. Francese, C. Gravino, M. Risi, G. Scanniello, and G. Tortora, "On the Use of Requirements Measures to Predict Software Project and Product Measures in the Context of Android Mobile Apps: A Preliminary Study," in *Proceedings of the 41st Euromicro Conference on Software Engineering and Advanced Applications (EUROMICRO-SEAA)*, pp. 357–364, 2015.

[3] P. C. Blumenfeld, E. Soloway, R. W. Marx, J. S. Krajcik, M. Guzdial, and A. Palincsar, "Motivating Project-Based Learning: Sustaining the Doing, Supporting the Learning," *Educational psychologist*, vol. 26, no. 3-4, pp. 369–398, 1991.

[4] J. R. Savery and T. M. Duffy, "Problem Based Learning: An Instructional Model and its Constructivist Framework," *Educational technology*, vol. 35, no. 5, pp. 31–38, 1995.

[5] L. K. Michaelsen, A. B. Knight, and L. D. Fink, *Team-Based Learning: A Transformative Use of Small Groups*. Greenwood publishing group, 2002.

[6] L. Vygotsky, *Thought and Language*. MIT Press, 2012.

[7] M. Bieliková, "Experiences with Designing a Team Project Module for Teaching Teamwork to Students," *Journal of computing and information technology*, vol. 13, no. 1, pp. 1–10, 2005.

[8] W. Humphrey and J. Over, *Leadership, teamwork, and trust: Building a competitive software capability*. Addison-Wesley Professional, 2010.

[9] M. E. Joorabchi, A. Mesbah, and P. Kruchten, "Real Challenges in Mobile App Development," in *Proceedings ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*, pp. 15–24, ACM Press, 2013.

[10] S. Kizaki, Y. Tahara, and A. Ohsuga, "Software Development PBL Focusing on Communication Using Scrum," in *Proceedeings of the IIAI 3rd International Conference on Advanced Applied Informatics (IIAIAAI)*, pp. 662–669, 2014.

[11] B. Silva and P. Girão, "A Study and a Proposal of a Collaborative and Competitive Learning Methodology," in *Proceedings of the IEEE Education Engineering (EDUCON)*, pp. 1011–1018, 2010.

[12] L. M. Regueras, E. Verdú, M. J. Verdú, and J. P. De Castro, "Design of a Competitive and Collaborative Learning Strategy in a Communication Networks Course," *IEEE Transactions on Education*, vol. 54, no. 2, pp. 302–307, 2011.

[13] J. C. Burguillo, "Using Game Theory and Competition-Based Learning to Stimulate Student Motivation and Performance," *Computers & Education*, vol. 55, no. 2, pp. 566–575, 2010.

[14] S. Yadav and J. Xiahou, "Integrated Project Based Learning in Software Engineering Education," in *Proceedings of the International Conference on Educational and Network Technology (ICENT)*, pp. 34–36, 2010.

[15] R. Francese, C. Gravino, M. Risi, G. Scanniello, and G. Tortora, "Using Project-Based-Learning in a Mobile Application Development Course - An Experience Report," *Journal of Visual Language and Computing*, vol. 31, pp. 196–205, 2015.

[16] L. Briand and I. Wieczorek, *Resource Estimation in Software Engineering*. Encyclopedia of Software Engineering, John Wiley & Sons, Inc. All, 2002.

[17] B. A. Kitchenham, E. Mendes, and G. H. Travassos, "Cross versus Within-Company Cost Estimation Studies: A Systematic Review," *IEEE Transaction on Software Engineering*, vol. 33, no. 5, pp. 316–329, 2007.

[18] M. J. Shepperd and G. F. Kadoda, "Comparing Software Prediction Techniques Using Simulation," *IEEE Transaction on Software Engineering*, vol. 27, no. 11, pp. 1014–1022, 2001.

[19] E. Mendes and N. Mosley, "Bayesian Network Models for Web Effort Prediction: A Comparative Study," *IEEE Transaction on Software Engineering*, vol. 34, no. 6, pp. 723–737, 2008.

[20] J. Ceddia and M. Dick, "Automating the Estimation of Project Size from Software Design Tools using Modified Function Points," in *Proceedings of the Sixth Australasian Conference on Computing Education*, vol. 30, pp. 33–39, Australian Computer Society, Inc., 2004.

[21] B. Bruegge and A. H. Dutoit, *Object-Oriented Software Engineering: Using UML, Patterns and Java, 2nd edition*. Prentice-Hall, 2003.

[22] C. Harpham and R. L. Wilby, "Multi-site Downscaling of Heavy Daily Precipitation Occurrence and Amounts," *Journal of Hydrology*, vol. 312, no. 1, pp. 235–255, 2005.

[23] T. J. Hastie and D. Pregibon, *Generalized Linear Models*. Wadsworth and Brooks/Cole, 1992.

[24] V. R. Basili, L. C. Briand, and W. L. Melo, "A Validation of Object-Oriented Design Metrics as Quality Indicators," *IEEE Transaction on Software Engineering*, vol. 22, no. 10, pp. 751–761, 1996.

[25] T. Gyimóthy, R. Ferenc, and I. Siket, "Empirical Validation of Object-Oriented Metrics on Open Source Software for Fault Prediction," *IEEE Transaction on Software Engineering*, vol. 31, no. 10, pp. 897–910, 2005.

[26] G. Scanniello, C. Gravino, A. Marcus, and T. Menzies, "Class Level Fault Prediction Using Software Clustering," in *Proceedings of the International Conference on Automated Software Engineering (ASE)*, pp. 640–645, IEEE Computer Society, 2013.

[27] S. S. Shapiro and M. B. Wilk, "An Analysis of Variance Test for Normality," *Biometrika*, vol. 3, no. 52, 1965.

[28] E. Mendes and B. Kitchenham, "Further Comparison of Cross-company and Within-company Effort Estimation Models for Web Applications," in *Proceedings of the International Software Metrics Symposium*, pp. 348–357, IEEE press, 2004.

Towards Formal Multimodal Analysis of Emotions for Affective Computing

Mehdi Ghayoumi
Artificial Intelligence Lab
Department of Computer Science
Kent State University
OH, USA
mghayoum@kent.edu

Maha Thafa
Artificial Intelligence Lab
Department of Computer Science
Kent State University
OH, USA
mthafar@kent.edu

Arvind K. Bansal
Artificial Intelligence Lab
Department of Computer Science
Kent State University
OH, USA
akbansal@kent.edu

***Abstract*—Social robotics is related to the robotic systems and human interaction. Social robots have applications in elderly care, health care, home care, customer service and reception in industrial settings. Human-Robot Interaction (HRI) requires better understanding of human emotion. There are few multimodal fusion systems that integrate limited amount of facial expression, speech and gesture analysis. In this paper, we describe the implementation of a semantic algebra based formal model that integrates six basic facial expressions, speech phrases and gesture trajectories. The system is capable of real-time interaction. We used the decision level fusion approach for integration and the prototype system has been implemented using Matlab.**

Keywords- Affective computing, Emotion recognition, Human-machine interaction, Multimedia, Multimodal, Decision level fusion, Social robotics.

I. INTRODUCTION

In the Human Compute Interaction (HCI) researches and studies, facial expression, speech and body movements have the major roles [8, 18 and 24]. Due to the aging society and increasing cost of health care, elderly care and assisted living, there has been significant interest in the development of social robotic systems that can interact with humans through the use of sensors. The social-robotic system can be humanoid, computers or intelligent machines as in Internet of Things who will interact with humans in the daily life. Interacting with humans requires understanding emotions [6] and emotions are based on a person's state of mind and partially regulated by personality, context and conditioning. Emotion is a language for communicating by feelings and it includes approval and disapproval to robotic systems. Interactive emotions [8, 16, and 18] are a subclass of human emotion analysis that humans use to interact with each other in close proximity. There are many interactive emotions that a person can express to machine during interaction, such as happiness, anger, embarrassment, surprise, rage, disappointment, confusion, elation, depression, approval and disapproval. Interactive emotions are expressed using a combination of verbal and nonverbal modes such as facial expressions [7, 10,], speech [17] including silence, gesture including body-posture and body-motion. Single mode may not give the emotion completely, or may be unavailable during emotive interaction. For example, the face may be occluded by the hands during sadness when a person is in deep pain or is crying. A person in shock or deep pain may not utter a single word. In the early stages of social robotics, most of the human-computer interaction in the service industry will involve brief commands or query by the human, and the robots will play a subordinate role rather than as companion role. Most of the emotion recognition will be limited to the integration of:

1) Facial expressions,
2) The limited amount of speech commands and emotional phrases to provide as an approval, disapproval, encouragement of a robot response or action, and
3) Gesture analysis of the upper body part involving head, hand, fingers, eyes, and lips.

The speech commands may be restricted to commands like "yes", "no", "don't like", "very good", "I am happy" etc. Some of these commands may have limited speech attributes such loudness showing disapproval or anger. Speech analysis can be done by a combination of text-to-speech conversion to understand the emotional phrases, and fast Fourier transform can be used to derive the variation of speech features such as energy, amplitude envelope, pitch during emotional variation. Real-time facial expression analysis in a video analysis where emotions and emotion transition can be studied by frame-to-frame analysis of facial expressions. Gesture analysis is done by analyzing video analysis and depth analysis as in Kinect based systems.

Currently, computational systems are limited to analyzing a single mode of emotion expression such as facial expression, speech, and (to some limited extent) multimodal analysis [22].

DOI reference number: 10.18293/DMS2016-030

The current integration of multimodal analysis systems of interactive emotions, lack:

1) A formal model to combine multiple modes such as facial expressions, speech analysis and gestures,
2) Complete catalog of upper body gestures, and
3) Capability of real-time analysis of facial-expressions and gestures.

In biometric systems, multimodal systems have some advantages which make the system accuracy and performance higher. Here we are using this model to achieve better results [23]. This research effort is in the direction of real-time integration of multimodal analysis system to derive emotion during HRI. A fusion module combines the weighted scores derived from each mode to derive the best emotion. The major contribution, here are:

1) Implementation of a real-time facial expression system based upon integration of geometric features, facial action units and facial symmetry [27],
2) Gesture recognition systems using fuzzy values,
3) Emotional Phrase lookup module,
4) Weighted score based Integration of a multimodal system based upon an abstract model of multimodal emotion analysis.

The rest of the paper is organized as follows. Section 2 describes background. Section 3 describes the overall architecture. Section 4 explains the speech recognition system and facial expression analysis will be explained in the section 5. Section 6 demonstrates the gesture modeling. Section 6 illustrates the implementation and performance results. Section 7 demonstrates the related works and the last section concludes the work, and describes the future directions.

II. BACKGROUND

This section describes the background material related to facial expression, speech analysis and gesture recognition and describe basic mathematical concepts needed for abstract modeling of the emotions.

A. Components of Emotion Recognition

There are three popular psychological theories of emotions: *James-Lange theory* [19], *Cannon-bard theory* [20] and *Schater-singer* theory [21]. *James-Lange* theory states that the mental state in response to the reactions which caused by external stimuli is emotion. *Cannon-Bard* theory is based upon anticipation rather than as a reaction to specific action. *Schachter-Singer* theory states that encountering an emotion requires both an interpretation of the bodily response as well as specific circumstance at a specific moment.

Also, there are three major classes of emotions:

1) Basic emotions,

2) Emotions that having same basic class, but having different intensity,

3) Mixed emotions that are a combination of one or more basic and/or mixed emotions.

Although, there are some disagreements among researchers, and a popular computational theory of Ekman [22] identifies six basic emotions: *happiness, sadness, surprise, disgust, anger* and *fear*. An example set of emotions having same basic class, but different intensities is {relaxed, happy, delighted, and euphoric}. Another set is {upset, anger, rage} etc. An example of mixed emotion is {amazed} that is a combination of {surprise and happiness} or {envy} which is the combination of {sadness and anger} or {despair} which is the combination of {fear and sadness}. In general, Facial Expressions have been done using these types of systems:

1) Facial Action Coding System (FACS) based on the simulation of facial muscle movement,
2) Geometric Features Modeling (GFM) based upon the movement of major feature-points of the face such as dynamic change in location endpoints and curvature of the mouth, eye, lips, forehead furrows and space between eyebrows.

The unit of FACS is an Action Unit (AU) that involves a segment of a muscle in facial expression. There are 17 major AUs involved in basic facial expressions. Examples of AUs involved in facial expressions are: inner brow raiser, outer brow raiser, brow lowered and drawn together, upper eye-lid raised, cheek raised, upper lip raised, lip corners pulled down, etc. The major geometric feature points, involved in facial expression analysis are given in Figure 1 which these features-points include:

1) 3 eyebrow points in each of the eyebrows:
$$b_1^L,\ b_2^L,\ b_3^L,\ b_1^R,\ b_2^R,\ b_3^R$$
2) 2 endpoints of eyes in each of the eyes:
$$e_1^L,\ e_2^L,\ e_1^R,\ e_2^R$$
3) Middle points eye-lid in each of the eyes:
$$el_L \text{ and } el_R$$
4) 2 endpoints of nose:
$$n^T \text{ and } n^B$$
5) 2 endpoints of mouth:
$$m^L \text{ and } m^R$$
6) 2 middle points of the mouth based on top and bottom lips:
$$m^T \text{ and } m^B$$
7) Chin-point denoted as:
$$ch.$$

The points shaded in dark black- e_1^L, e_2^L, e_1^R, e_2^R, n^T and n^B do not move, and act as reference-points. Remaining spotted-points move with emotions, and their displacement is used to derive the facial expression.

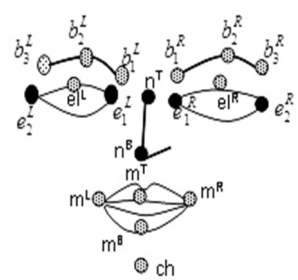


Figure 1. Major feature-points on the face

Emotional speech has multiple features such as phonemes, emotional phrases, amplitude, syllable envelope, pitch, rhythm, quantile and silence. Phonemes are the basic units of speech. During emotional interaction, pitch, amplitude, syllable envelope, duration of silence and utterances change significantly, act as parameters for the recognition of interactive emotions. Gesture is a nonverbal communication using perceptible bodily actions such as body-postures and body-part movements, including movements of the head, torso, hands, face and eyes. Different components of the emotions are measured using different sensors. Facial-Expression uses image analysis techniques to identify the movement of facial feature points, speech analysis uses wavelet analysis, FFT analysis, morphology analysis, text-to-speech conversion for phoneme detection and dictionary lookup to identify phrases. Gesture recognition requires image analysis to derive postures and video-frame analysis to derive motion of various body parts such as head, arm, eyes, hand, palm, fingers. The posture and motion are modeled as fuzzy values to reduce the computational space. The motion of the body parts can also be derived using skeletal and depth analysis used in Kinect.

B. Mathematical concepts

The Fuzzy values map a large value-space to a smaller finite space. The major advantages of the use of fuzzy values are:

1) Reduction of the computational complexity
2) Nearness to human perception and
3) Tolerance from the sensor noise.

We use two types of fuzzy sets:

1) Discrete fuzzy set, and
2) Ordered fuzzy sets.

A discrete fuzzy set has values that have no relationship that shows transitivity. For example, a head posture can be {rotated-left, rotated-right, normal, tilted-left, tilted-right, looking-down, looking-up}. An ordered fuzzy set shows transitive relationship between the values, and is used to model motion intensity in gesture analysis for better classification of emotion. For example, the speed of a head-motion can be modeled as {still, slow, normal, fast, very fast}. The values in the fuzzy set can be mapped onto the ordinals 0… 4:

Still →0, Slow→1, Normal→2, Fast→3 and Very Fast→ 4 (1)

The use of this mapping allows the use of comparison operators on ordered fuzzy sets. Cartesian product of the N sets returns a set of N-tuples such i^{th}-field of an element is a member of the ith set as shown:

$$X_1 \times \ldots \times X_n = \{(x_1, \ldots, x_n) \mid x_i \in X_i \forall i = 1, \ldots, n\} \quad (2)$$

Two domains can be joined using:

1) Product-domain that uses the Cartesian product $A \times B$, or

2) A sum - domain that uses disjoint-union $A + B$, or

3) *Function Domain* mapping on lifted domains f: A⊥ B.
Where ⊥ is the bottom symbol used to catch all ill-defined mappings.

III. Overall Architecture

Overall architecture (see Figure 2) has six major modules:

Unit 1 - Facial Expressions Subunit (FE)

The subunit takes a video-clip that is a sequence of frames, and extends the integration of FACS + geometric feature analysis technique for real-time basic face-expression recognition to derive the ranked subset of facial expressions for each frame in the video-clip. The analysis of a frame may result in more than one facial expression due to the:

1) Partial or full occlusion of the face due to gesture or head rotation,
2) Transition of emotions,
3) Inherent accuracies in the facial expression technique,
4) A variation of the facial expression due to the situation, personality or culture,
5) Low emotional intensity.

The outcome of this facial-expression analysis gives a sequence of subsets of derived possible emotions with the matching score of the form:

$$<FE_1, \ldots, FE_j, FE_{j+1}, \ldots, FE_N> \quad (3)$$

Where N is the number of frames in the clip, FE_i is a subset of rank facial expressions of the form

$$\{(e_1, s_1) \ldots (e_m, s_k)\} \text{ For } k \geq 1, e_i \in \Sigma, s_i > \text{threshold and } s_i > s_{(i+1)} \quad (4)$$

Which e and s are the emotion and its corresponding score respectively.

Unit 2 Gesture Fuzzy Parameterization Module (GFP)

The model measures different postures and motion intensity and frequency of different emotional gesture patterns.

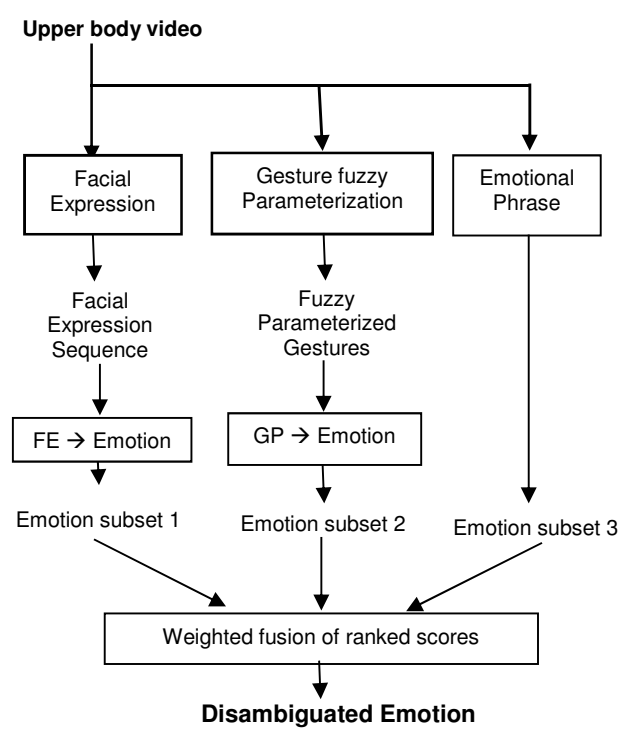


Figure 2. Overall architecture of multimodal fusion

The different postures are: body-posture, head-posture, shoulder-posture, hand-postures, palm-posture, finger-postures and eye-gaze and the various-motions are: head-motion, arm-motions, eye-motions, finger-motions. The details of the fuzzy parameterization and emotional head-motion gesture is given in Section 5.

Unit 3 - Emotional Phrases Module

Emotional phrase modules use a hash function to generate the index, and stores multiple emotions with a fuzzy intensity value. Once a phrase is recorded, then text-to-speech converter is used to derive the text. Individual words are looked up in the user-specified dictionary to remove the noise in text-to-speech conversion. The speech analysis system is used to derive the relative energy level. The energy level is parameterized to a fuzzy value, and the hash function is used in the derived text to identify the index of the speech. Using this index, the corresponding set of emotions that closely matches the intensity levels are derived.

Unit 4- FE (Emotion Module)

Many adjacent frames will have the same subset of facial expressions until the facial expression changes. We call this frame as the emotion transition point. Thus:

$$<FE_1, \ldots, FE_j, FE_{j+1}, \ldots, FE_N> \rightarrow <(E_1, d_1), (E_2, d_2)\ldots, (E_M, d_M)> \quad (5)$$

Where E_I $(1 < I < M)$ is a subset of Σ, and d_I is of the form start-frame: end-frame. The term d_I is used:

1) To reconcile the emotion in the fusion-module, and
2) To derive the duration of emotion to match with the duration of emotion derived from gesture analysis and emotional phrase analysis.

Unit 5- GFP (Emotion Module)

This module takes the fuzzy parameterized values of different body parts and their motion, concatenates them into a long string, and creates a string, and performs a similarity-based search in K-dimensional space where K is the number of fuzzy-valued component to identify a possible set of emotions. The attributes of fuzzy-vector representation of head-motion trajectory is hashed to derive the possible set of emotions based upon gesture. The duration of the body-motion is noted like unit 4, and its output is also on the form:

$$<(E_1, d_1), (E_2, d_2)\ldots, (E_M, d_M)> \quad (6)$$

Where E_I $(1 < I < M)$ is a subset of Σ, and d_I is of the form start-frame: end-frame.

Unit 6 - Weighted Fusion of Emotion Module

The role of the weighted fusion module is to:

1) Fuse the information of the ranked emotions from the three modules to reduce ambiguity and improve ranking scores,
2) Derive the duration of emotion.

The input of the FE (emotion module (unit 4), and GP (emotion module (unit 5) and unit 3 are a sequence of set of rank emotion with the duration. Under the assumption, that emotions are expressed involuntarily in the facial expressions first, the start frame of the facial expression should occur before followed by the emotions derived from other two modules. To handle the issues that emotions may not be expressed by one or more modules, the weight is dependent upon:

1) Availability of the emotions from the specific mode,
2) The noise level.

For example, initial weight is w_1, w_2 and w_3 for the fusion of the corresponding modes, the weights w_k $(1 \leq i \leq 3)$ is altered by:

$$\left(\sum_{i=1}^{i=3} w_i \Big/ \sum_{i \neq j, i=1}^{i=3} w_i\right) \quad (7)$$

Since one of the modes is missing. The fusion is performed by:

1) Multiplying the ranked score of each emotion by the corresponding weight,
2) Adding the scores of the same derived emotions from different weights, and
3) Sorting the emotions by the descending order of the scores, and picking up the emotion with the highest score.

If the top two scores are very close, then it can be a case of emotional transition or mixed emotion.

IV. FACIAL EXPRESSION ANALYSIS

We extend the integration of FACS system interaction and geometric feature analysis [8] to make the facial expression analysis by using facial symmetry and invariance under head-motion. There are 13 moving-points (11 active points and 2 passive points) and 6 references-points. FACS system analysis has been used to derive the features-points that are significant during the expression of a specific facial expression. For example, for a surprise the all eyebrow points are uniformly raised; for happiness mouth corners are stretched, the eye-lid point gets lowered; for anger distance between eyebrows becomes smaller, inner eyebrow points get lowered. These FAUs have been translated to the corresponding feature-point movements as given in Table 1. We denote vertical-up motion by ↑, vertical-down motion by ↓, horizontally stretched outwards by '↔', horizontally compressed inwards by '↢', oblique-stretched downwards by '↘', oblique-stretched upwards by '↗'. If the emotion is symmetric, then the subscripts L and R have been omitted. If the movement is optional or shows higher intensity increase then it has been placed within the square brackets. Conjunction has been shown using concatenation. Essential feature-point have been within parenthesis () separated by ','. At least one of the essential feature point motion has to be present for the emotion to occur. Scores are associated with the presence of each feature-point motion.

TABLE1. Feature Point displacements

Facial Expressions	Major Feature-points displacements
Anger	$(e_1$↢ el↑$) + [e_2$↑$] +$ $[m^T$↑ m^B↑$]$
Disgusted	$(m^T$↑ch↑$)$ $+ [\{m^L, m^R\}$↓$] + [m^B$↑$]$
Fear	$(e_1$↑$, m^L$↓ m^R↓$) + [m^T$↓$] +$ [e1↢]
Happiness	$(m^L$↗m^R↗$, M^T$↑m^B↓ ch↓m^L↔m^R↔$)$
Sadness	(el↓ m^L↔m^R↔$) +$ [ch↓]
Surprised	$(e^1$↑e^2↑e^3↑el↑ch↓$) + [m^T$↑m^B↓$]$

For each feature point, we measure the displacement distance and the direction of the displacement. Thus the derivable facial expressions are mapped to a vector of (displacement-distance ratio, direction).

Direction is a discrete-fuzzy set with six possible values:

1) Vertical-up,
2) Vertical-down,
3) Horizontal-compressed-inwards,
4) Horizontal-stretched-outwards,
5) Oblique-stretched-upwards, and
6) Oblique-stretched-downwards.

Facial expressions can be occluded due to:

1) Gestures such as hand covering face in case of sadness,
2) Lighting conditions, and
3) Head rotation or tilt.

In order to complete the information, we use the facial symmetry around nose to fill in the information about those facial expressions that show symmetry such as happiness, anger, surprise, sadness and fear. Disgust may shows asymmetric features. In order to variation of the displacement projection due to head motion, we use the distance-ratio (point-displacement from the relaxed state / distance between reference-points) which go with similar transformation.

For example, to keep the horizontal displacement we use distance between the two outer-eye corners: e_2^L and e_2^R in the denominator of the ratio; and for the two noses displacement. We use the distance between the two noses-points and n^B.

There are two types of motion-points:

A. Points that move in only in vertical up-down direction, such as:

$$e_1^L, e_2^L, e_1^R, e_2^R \text{ and}$$

B. Points that move in all four directions: vertical up-down and horizontal inside-outside motion such as:

$$el^L, el^R, m^L, m^R, m^T, m^B, ch$$

For the points that show motion in vertical up-down direction have only one entry in the *facial-expression ratio vector*, and points that show motion in up-down and stretch-compression mode have two entries in the vector. Based upon the emotion, some of the entries may not change during that emotion.

For example, in surprise, only, $el^R, e_2^L, e_3^L, e_2^R, e_3^R, m^T, m^B$ and ch change. This characteristic of facial-expression ratio-vector provides invariance against head-motion as well as specific characterization of facial-expressions.

V. Speech Analysis

Embedding the component of emotion processing into existing speech systems makes them more natural and effective. Several approaches to recognize emotions from speech have been reported. In a conversation, non-verbal communication carries an important information like the intention of the speaker. In addition to the message conveyed through text, the manner in which the words are spoken, conveys essential non-linguistic information. The same textual message would be conveyed with different semantics by incorporating appropriate emotions. Spoken text may have several interpretations, depending on how it is said. For example, the word 'OKAY' in English, is used to express respect, disbelief, agreement, and disinterest. Therefore, understanding the text alone is not sufficient to interpret the semantics of a spoken utterance. However, it is important that, speech systems should be able to process the non-linguistic information such as emotions, along with the message. Choosing suitable features for developing any of the speech systems is a crucial decision.

We have three important speech features, namely:

1) Excitation source,
2) Vocal tract system, and
3) Prosodic features.

VI. Gesture Analysis

Gesture parameterization has two modules:

a) Deriving the posture of upper-body parts and their motions,

b) Mapping fuzzy to actual values to reduce the search space.

These fuzzy values are concatenated so that all the values from discrete sets are concatenated together, and all the values from the ordered sets are grouped together. This separation is necessary because mismatch in discrete sets leads to failure, while mismatch in ordered sets is permissible. Abstract modeling of emotion requires functional mapping of Cartesian product of different components to derivable emotions. Since all the component tuples may not map to valid emotional elements in the emotional domain, we make the emotion domain a lifted domain by introducing a bottom symbol $\perp$ in the set of well-defined emotions. The lifted domain allows for catching the error conditions when the tuple of fuzzy component values do not map to any specific emotion.

Fuzzy values are calculated using statistically derived thresholds. There are two types of sets:

1) Discrete sets where the values are not ordered and ordered sets,
2) Ordered sets are used in modeling the extent of posture variation and intensity of the motions of various gestures.

The parameters of the motion are:

a) Start-position,
b) End-position,
c) Frequency,
d) Speed,
e) Attack,
f) Relaxation.

The attack is the rate of change of speed until the motion attains the peak speed, relaxation is the rate of reduction of speed to the speed reduces from the peak speed to no motion. The mapping of the components is described by equation (7).

$$\text{Posture}_1 \times \ldots \times \text{Posture}_M \times \text{Motion}_1 \times \ldots \times \text{Motion}_M \rightarrow \text{set of possible emotions} \quad (7)$$

VII. Implementation and Performance Results

The implementation can be divided into 3 major steps as follows:

A. *Implementing Facial Expression Sequence Analysis*

We extract the main parts of the face such as eyes, eyebrow, nose and mouth, then find the key points in each segment as shown in the figure 1 then we extract the feature vectors from extracted key points and train the networks.

B. *Implementing Head Gesture Analysis*

Human head movement is very important in general conversation and communication. Despite the influential role of the head gestures, very little research has examined the gesture's role in the robot-human interaction process. In software module, the pose of the human head is estimated with a constraint that the human head is a 3 DOF rigid object which has yaw, pitch and roll movements. We have used geometric head pose estimation algorithm which estimates the head pose through a standard webcam of the computer. The details of this algorithm can be found in [12].

C. *Implementing Emotional Phrase Matching*

A cepstrum is obtained by computing the Fourier Transform of the logarithm of the spectrum of a signal. There are different kinds of cepstrum such as complex cepstrum, real cepstrum, phase cepstrum and power cepstrum. The power cepstrum is used in speech synthesis applications and here we use it. The approach based on decision-level fusion obtained. The performance of the classifier was 94.6%, both for the best probability and for the majority vote plus best probability approaches.

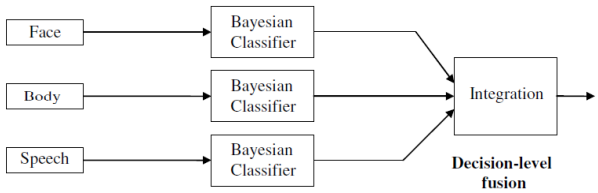


Figure 3. The decision level fusion

Table 2 shows the performance of the system with decision level integration using the best probability approach. Anger has the emotion recognized with highest accuracy.

TABLE 2. Decision level integration with the best probability approach

Anger	Happy	Sad	Surprise	Disgust	Fear	
98.3	0	0	0	4.3	0	**Anger**
0	**95.4**	0	7.2	0	0	**Happy**
3.1	0	**92.1**	0	2.7	2.2	**Sad**
9.3	10.4	0	**87.5**	2.3	5.8	**Surprise**
7.2	0	4.1	3.3	**90.2**	4.2	**Disgust**
0	0	0	12.2	11.1	**83.3**	**Fear**

VIII. Related Works

There are many related works in the facial expression analysis [1, 10, 25 and 26], gesture analysis [2, 4, 12, and 18] emotion recognition in speech [9] and multimodal fusion [3, 4]. Castellano et al. [7] extended their work to multimodal framework integrating face-expression, body-gestures, and speech. A Bayesian classifier was used for feature level fusion and decision level fusion. A comparison between unimodal, bimodal, and multimodal classification showed that multimodal classification is better. There is an additional need to identify features that are relevant to the dynamics of expressive emotions, however, their study is limited to identifying eight emotions [7]. We have been influenced by the research to derive from the research to analyze facial expressions based upon action units and map action units based movement to geometric feature-point movements [8]. The use of geometric feature-points and fusion allows for better accuracy in our research. In addition, we use abstract model for gesture analysis. Our geometric point movement based upon study of facial action units is generally enough to analyze finer classification of emotions and mixed emotions. Many interesting works about audio-visual fusion/mapping has been proposed for multimodal information processing. For instance, speech based facial animation [13], and audio-visual based emotion recognition [14]. There is also some work for head motions [15] and body gestures [16], however, most of them just focused on the gesture recognition.

IX. Conclusion and Future Works

In this paper, we have described a detailed methodology and an initial prototype implementation of real-time multimodal fusion to derive interactive emotion for interaction with social-robots and intelligent machines with limited emotional phrase based interaction. The proposed integrated system has many novelties such as: an abstract model of fusion based upon a semantic algebra that the maps Cartesian product of different components to derivable emotions, the use of invariant displacement of geometric feature-points to identify facial-expressions, and Gestures based upon head-trajectory and fuzzy values of other upper body parts to reduce the search space. Currently, the gesture based system is limited to, image analysis of feature-points in the head and hand to derive posture. We are looking into Kinect based analysis to integrate skeleton based body posture, motion and depth analysis [15] for better accuracy.

References

[1] R. Adolphs. "Recognizing emotion from facial expressions: psychological and neurological mechanisms," *Behav. Cogn. Neurosci. Rev.*, Vol. 1, 2002, pp. 21-62.

[2] V. Bevilacqua, D. Barone, F. Cipriani, G. D'Onghia et al., "A new tool for gestural action recognition to support decisions in emotional framework", *Proceedings of the IEEE Symposium of Innovations in Intelligent Systems and Applications* (INISTA), 2014, pp. 184-191.

[3] G. Caridakis, G. Castellano, L. Kessous, A. Raouzaiou, L. Malatesta, et al. "Multimodal emotion recognition from expressive faces, body gestures and speech"; *Artificial Intelligence and Innovations: From Theory to Applications*, Springer Berlin Heidelberg, 2007, pp. 375-388.

[4] G. Castellano, S. D. Villalba and A. Camurri. Recognizing human emotions from body movement and gesture dynamics"; *Affective Computing and Intelligent Interaction*, Springer Berlin Heidelberg, pp. 71-82.

[5] P. Ekman and W. Friesen, "Facial Action Coding System: A Technique for the Measurement of Facial Movement," *Consulting Psychologists Press*, Palo Alto, 1978.

[6] J. M. Fellous and M.A.Arbib,"Who needs emotions? The brain meets the robots", Oxford press, 2005.

[7] L. Gang, L. Xiao-hua, Z. Ji-Liu and G. Xiao-gang, "Geometric feature based facial expression recognition using multiclass support vector machines," *IEEE International Conference on Granular Computing (GRC '09)*, 2009, pp. 318-321.

[8] M. Ghayoumi and A. K. Bansal, "Unifying Geometric Features and Facial Action Units for Improved Performance of Facial Expression Analysis," Proceedings *of the International Conference on Circuits, Systems, Signal Processing, Communications and Computers* (CSSCC 15), pp. 259-266.

[9] C. H. Lim, E. Vats and C. S. Chan, "Fuzzy human motion analysis: A review," *Pattern Recognition,* Vol. 48, 2015, pp. 1773-1796.

[10] S. Mitra and T. Acharya. "Gesture recognition: A survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, Vol. 37, 2007, pp. 311-324.

[11] C. M. Lee, S. S. Narayanan, "Towards Detecting Emotions in Spoken Dialog," *IEEE Trans. On Speech and Audio Processing*, Vol. 13, No. 2, 2005, pp. 293-303.

[12]] M. Khan, S. Rehman, Z. Lu and H. Li, "Head Orientation Modeling: Geometric Head Pose Estimation using Monocular Camera," in Proceedings of the 1st IEEE/IIAE International Conference on Intelligent Systems and Image Processing, 2013.

[13] Alberto B, Piero C, Giuseppe RL, Giulio P Lucia" a new WebGL-based talking head," 15th Conference of the International Speech Communication Association, Singapore 2014.

[14] Cowie, R, Douglas-Cowie, E "Emotion recognition in human-computer interaction." IEEE Signal Processing Magazine. pp. 33-80, 2001.

[15] Bo X, Georgiou Panayiotis G, Brian Baucom, Shrikanth S Narayanan, "power-spectral analysis of head motion signal for behavioral modeling in human interaction," I.E. International Conference on Acoustics, Speech, and Signal Processing, 2014.

[16] Welbergen HV, Reidsma D, Ruttkay ZM, Zwiers EJ D, "A BML realizer for continuous, multimodal interaction with a virtual human." Multimodal User Interf 3D4:271-284, 2010.

[17] Oh-Wook Kwon, Kwokleung Chan, Jiucang Hao, Te-Won Lee, "Emotion Recognition by Speech Signals" - INTERSPEECH, 2003.

[18] M. Ghayoumi, A. Bansal, "An Integrated Approach for Efficient Analysis of Facial Expressions", SIGMAP 2014.

[19] A, Walter. "The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory." The American Journal of Psychology 39: 106-124.

[20] Friedman, B.H. "Feelings and the body: The Jamesian perspective on autonomic specificity of emotion." Biological Psychology 84: 383-393 (2010).

[21] Cotton, J. L. "A review of research on Schachter's theory of emotion and the misattribution of Arousal." European Journal of Social Psychology 11: 365-397 (1981).

[22] P. Ekman, Facial expression and emotion. American, 8 (4): 384-392, 199.

[23] M. Ghayoumi, "A Review of Multimodal Biometric Systems Fusion Methods and Its Applications.", ICIS, USA, 2015.

[24] H. Abrishami Moghaddam and M. Ghayoumi, "Facial Image Feature Extraction Using Support Vector Machines.", Proc. VISAPP, Setubal, Portugal, 2006.

[25] Ghayoumi, M., Bansal, A. K."Multimodal Architecture for Emotion in Robots Using Deep Learning." Future Technologies Conference, San Francisco, United States, FTC 2016.

[26] M. Ghayoumi, A. K. Bansal, "Emotion in Robots Using Convolutional Neural Networks.", ICSR 2016.

[27] M. Ghayoumi, A. K. Bansal, "Real Emotion Recognition Algorithm by Detecting Symmetry Patterns with Dihedral Group.", MCSI 2016.

Parameter Calibration Method of Microscopic Traffic Flow Simulation Models based on Orthogonal Genetic Algorithm

Yanfang Yang
School of Traffic and Transportation
Beijing Jiaotong University
Beijing, P. R. China
E-mail: yangyf@bjtu.edu.cn

Yong Qin
State Key Laboratory of Rail Traffic Control and Safety
Beijing Jiaotong University
Beijing, P. R. China
Corresponding author, Email: yqin@bjtu.edu.cn

Honghui Dong
School of Traffic and Transportation
Beijing Jiaotong University
Beijing, P. R. China
Email: hhdong@bjtu.edu.cn

Qing Zhang
Beijing Research Center of Urban Traffic Information Sensing and Service Technologies
Beijing Jiaotong University
Beijing, P. R. China
Email: 15125808@bjtu.edu.cn

***Abstract*—Traffic microscopic traffic simulation models have become extensively used in both transportation operations and management analyses, which are very useful in reflecting the dynamic nature of transportation system in a stochastic manner. As far as the microscopic traffic flow simulation users are concerned, the one of the major concerns would be the appropriate calibration of the simulation models. In this paper a parameter calibration method of microscopic traffic flow simulation models based on orthogonal genetic algorithm is presented. In order to improve the capacity of locating a possible solution in solution space, the proposed method incorporates the orthogonal experimental design method into the genetic algorithm. The proposed method is applied to an arterial section of Ronghua Road in Beijing. Through comparing with the parameter calibration method based on genetic algorithm, the advantage of the proposed method is shown.**

Keywords-Microscopic traffic flow simulation model; Parameter calibration; Orthogonal genetic algorithm; VISSIM

I. INTRODUCTION

Traffic simulation has become an important and popular tool in modeling transport system, with the progress of simulation technologies [1]. Traffic simulation models could be divided into three categories, including microscopic, macroscopic, mesoscopic simulation models. Microscopic simulation models simulate traffic at a level of individual vehicles [2]. Car-following and lane-changing models are the two fundamental components in microscopic simulation models. Macroscopic simulation models simulate transportation network section-by-section rather than tracking individual vehicles. Mesoscopic traffic simulation models combine the properties of the microscopic and macroscopic simulation models. For the traffic simulation models, the simulation results depend on the initial choice of the model [3] and the success of the calibration process [4].

The calibration of traffic microscopic simulation models is defined as the process of finding optimal parameters to match the field data so model will accurately represent field measure or observed traffic condition [5]. The optimization task involves comparing and minimization differences of selected indicators, e.g., travel time and queuing length [4], delays [6], travel time distribution [7], saturation flow rates [8] and emission [9], between the calibration model and the ones counted and measured in local traffic network.

Calibrating traffic model of bigger special and time scopes of a traffic network needs deal with a larger number of input parameters and calculating processes. In order to decrease time consumption, artificial intelligent techniques are applied into the calibration of traffic microscopic simulation model. Genetic algorithm (GA) has become the most common used calibration algorithm for input parameters of the simulations [5, 10-13], since Cheu et al., firstly used GA calibrating FRESIM model [14]. Other intelligent algorithms are also used in the calibration of traffic simulation, e.g., perturbation stochastic approximation (SPSA) scheme [15], particle swarm optimization (PSO) [16], and neural network approach [4]. These methods automate the calibration process to a certain degree and it was generally reported that they improve simulation performance over the default model parameter values.

Microscopic simulation is a complex system that all parameters work together to influence its modeling results. In calibrating such a complex model, users could get trapped in

DOI reference number: 10.18293/DMS2016-047

the local optima of the objective function, due to the high dimension and numerous local optima. This paper focuses on the above question of GA when calibrating the driving behavior model parameters in VISSIM.

The exposition of this paper is as follows: the next section implements the orthogonal genetic algorithm (OGA). Some studies have found that applying an experimental design method (orthogonal design) into GA may overcome the limitation mentioned above [17-18]. The third section gives the procedure of calibrating the microscopic traffic flow simulation model based on OGA. The fourth section applies OGA calibration method to the calibration of a signal intersection in Beijing, and comparing with the GA and orthogonal design method respectively. The final section summarizes the paper.

II. CALIBRATION METHOD BASED ON OGA

The proposed calibration method employs an orthogonal genetic algorithm. This section contains a brief overview of the VISSIM calibration parameters set, and the fundamentals of the OGA, including the structure of chromosome, fitness function and orthogonal crossover decoding.

A. Selection of Parameters

In this paper, a microscopic traffic flow simulation, VISSIM, is selected as the basic platform for the parameter calibration. VISSIM models the psychophysical driver behavior and attempts to capture both the physical and the human components of traffic [19]. Parameters of two driving behavior models are in considered in this paper: the car-following model and the lane-changing model. After parameter sensitivity analysis, i.e., one-way Analysis of Variance (ANOVA)[20], four parameters are selected as the calibration parameters. Table 1 lists the calibration parameter set, including default value, the minimum and maximum value.

TABLE 1 CALIBRATION PARAMETER SET

Parameters (x_j)	Unit	Default Value	Min (u_j)	Max (v_j)
x_1 Average standstill distance	m	2	0.5	3
x_2 Additive part of desired safety distance	NA	2	0.5	3
x_3 Multiple part of desired safety distance	NA	3	1	6
x_4 Maximum deceleration	m/s^2	4	2	6

B. The Structure of Chromosomes

Supposing P_i (x_{i1}, …, x_{iN}) is the ith chromosome, x_{ij} is the jth parameter value in the ith chromosome. $\forall$ i=1,2, …, M, j=1, 2,…, N=4. M is the total number of chromosomes and N is the number of parameters to be calibrated. [$\boldsymbol{l}$, $\boldsymbol{u}$]= [(l_1, …, l_N), (u_1, …, u_N)] defines the feasible solution space and the corresponding domain of x_j is [l_j, u_j] (e.g., the domain of x_1 is [0.5, 3] as listed in Table 1). In this paper, parameters to be calibrated are coded into chromosomes, which quantized by orthogonal design. We quantize the domain [l_j, u_j] of x_j into Q levels, where the design parameter Q is odd.

Algorithm 1 shows the procedure of constructing chromosome. Firstly, we calculate M, where $M= Q^J$, J is the smallest positive integer fulfilling $J \geq \frac{\log(N(Q-1)+1)}{\log Q}$.

Secondly, we construct the orthogonal array $L_M(Q^N)$ corresponding to the chromosome. Each element a_{ij} of the orthogonal array $L_M(Q^N)$ represents the levels numbers in orthogonal design, $a_{ij} \in \{0,1,...,Q\text{-}1\}$, $\forall$ i=1,2, …, M, j=1, 2,…, N. Finally, the corresponding parameter value x_{ij} of a_{ij} in feasible solution pace [l_j, u_j] is calculated by the equation as shown in follows:

$$x_{ij} = l_j + a_{ij} \times \frac{u_j - l_j}{Q-1}, \quad \forall 0 \leq a_{ij} \leq Q-1 \qquad (1)$$

Algorithm 1: Constructing chromosomes

Step 1. Calculating the number of chromosome, M=Q^J. J is the smallest positive integer fulfilling $J \geq \frac{\log(N(Q-1)+1)}{\log Q}$.

Step 2. Construct the orthogonal array $L_M(Q^N)$

(1) Construct the basic columns:

for k=1 **to** J **do**

$$j = \frac{Q^{k-1}-1}{Q-1}+1$$

for i=1 **to** Q^J **do**

$$a_{ij} = \left\lfloor \frac{i-1}{Q^{J-1}} \right\rfloor \textbf{ mod } Q$$

end for

end for

(2) Construct the non-basic columns:

for k=2 **to** J **do**

$$j = \frac{Q^{k-1}-1}{Q-1}+1$$

for s=1 **to** j-1 **do**

for t=1 **to** Q-1 **do**

$$a_{j+(s-1)(q-1)+t} = (a_s \times t + a_j) \textbf{ mod } Q$$

end for

end for

end for

(3) Selecting the first N columns to construct the chromosome encoding array $L_M(Q^N)$

Step 3. Calculate x_{ij} to construct chromosome using (1).

C. Fitness Function

The fitness function is a combination of the root mean absolute square error (RMASE) of travel time (TT) and maximum queue length (MQL) and between the VISSIM output and field data. The fitness function takes the form of

$$F_{total} = \frac{1}{\xi \cdot E_1 + (1-\xi)E_2} \tag{2}$$

$$E_n = \sqrt{\frac{1}{T}\sum_t \left(\frac{d_{nt}^s - d_{nt}^o}{d_{nt}^o}\right)^2} \quad \forall \quad n = 1,2\,, t = 1,2,\ldots,T \tag{3}$$

where
E_1 = the RMASE of TT
E_2 = the RMASE of MQL
T= the number of detector station;
d_{nt}^s = the simulation model output of tth sensor
d_{nt}^o = the field data of tth sensor station.

D. Adaptive orthogonal crossover operator

For each pair of parents (denoted by $P_1(x_{11}, x_{12}, x_{13}, x_{14})$ and $P_2(x_{21}, x_{22}, x_{23}, x_{24})$), performing the adaptive orthogonal crossover operation with the probability of crossover p_c. Adaptive orthogonal crossover algorithm is shown in Algorithm 2. σ_0 is the similarity threshold of each dimension between two parents .

Algorithm 2: Constructing chromosomes

Step 1. Calculating the number of similar dimension using $b = Num(|x_{1,j} - x_{2,j}| > \sigma_0)$

Step 2. Constructing orthogonal array $L_E(F^b) = [a_{ts}]_{E\times b}$, with b factors and F levels, using Algorithm 1.

Step 3. Generating E chromosomes as the potential offspring, $P_t'(x_{t,1}', x_{t,2}', x_{t,3}', x_{t,4}')$, $\forall$ t=1,2,…,E.

for t=1 **to** E **do**
  s=1
  for j =1 **to** 4 **do**
    if $|x_{1j} - x_{2j}| \le \sigma_0$
      $x_{tj}' = \dfrac{x_{1j} + x_{2j}}{2}$
    else if $|x_{1j} - x_{2j}| > \sigma_0$
      $x_{tj}' = \min(x_{1i}, x_{2i}) + a_{ts}\dfrac{\max(x_{1i}, x_{2i}) - \min(x_{1i}, x_{2i})}{F-1}$
    s= s+1
    end if
  end for
end for

Step 4. Calculating the fitness values corresponding to each potential offspring and selecting the chromosome corresponding to the max fitness value as the offspring of two parents P_1 and P_2.

In this paper， σ_0 =0.005 and F=2.

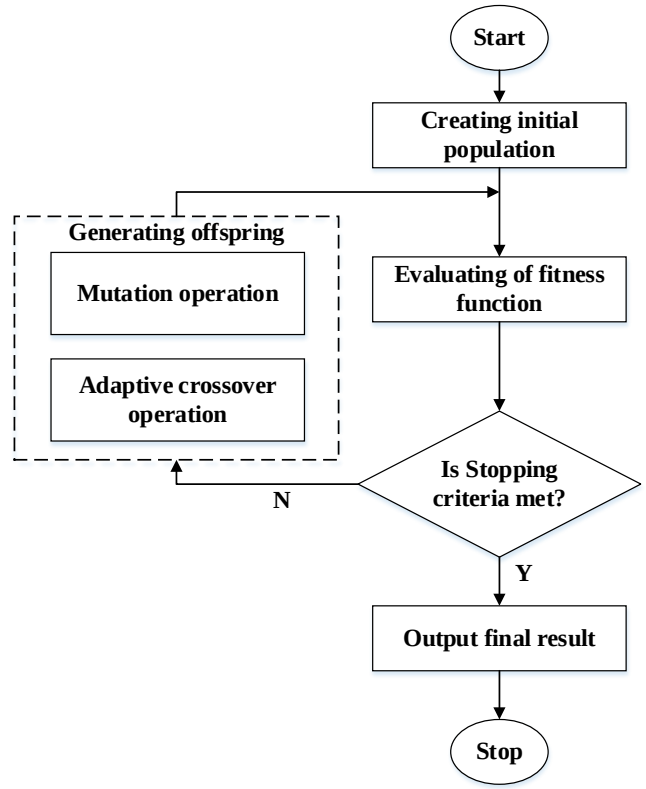


Figure 1 Calibration procedure

III. CALIBRATION PROCEDURE

We develop the parameter calibration program using VB language. The details of overall algorithm are as follows (see Fig. 1).

Step 1：Initialization

Executing Algorithm 1 to generate M potential parents (e.g., levels Q = 9 and $M = Q^J = 9^2 = 81$). The microscopic simulation traffic simulation model VISSIM is run with the M parameter group as the input file successively. Then M fitness values corresponding to the potential parents are calculated. Sort the potential parents in a descending sort order, according the fitness values. Select the first I := 50 parents as the potential initial population P_0. In order to keep the individual distributing uniformly, we select D := 10 chromosomes randomly from the D potential initial population as the initial population, denoted by $P_{gen=1}$.

Step 2：Adaptive crossover operation

For an arbitrary pair of parents, adaptive crossover operation is performed with the probability of crossover P_c (e.g., P_c =0.75) using algorithm in Algorithm 2 and generate offspring population C_{gen}.

Step 3：Mutation operation

Each chromosome in P_{gen} would undergo mutation operation with the probability of mutation p_m (e.g., p_m = 0.1). The mutation operation is as follows: (1) randomly generating an integer $j \in$ [1, N] and a real number $z \in [l_j, u_j]$; (2) replacing

the jth component of the chosen chromosome by z to get a new chromosome. The mutation operation generates a new population denoted by G_{gen}. The fitness values corresponding to each new chromosome in G_{gen} are calculated by running VISSIM model.

Step 4: Selection operation

In order to maintain the population diversity, we sort the population (P_{gen} + C_{gen} + G_{gen}) in a descending sort order according the fitness values , then select the first $\lfloor D*70\% \rfloor$ chromosomes and randomly select $D-\lfloor D*70\% \rfloor$ chromosomes from the rest of (P_{gen} + C_{gen} + G_{gen}) to constructing the next population P_{gen+1}.

Step 5: Check stopping rules

Supposing the max fitness values of iteration gen is $F_{\max}^{gen}$. If gen = the maximum number of iterations or $|F_{\max}^{gen}-F_{\max}^{gen-1}|\leq 0.005$, the program stops. Otherwise, go to Step 2 and gen = gen+1.

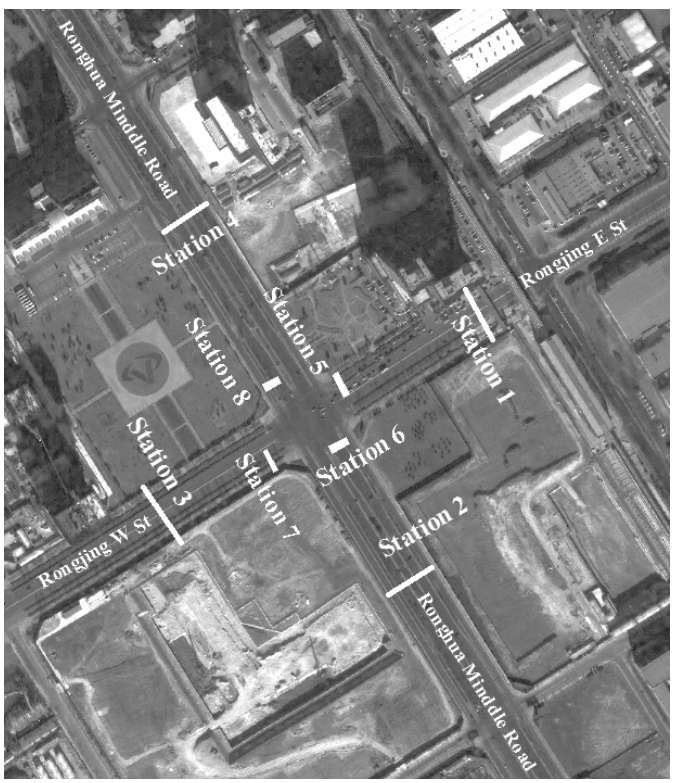


Figure 2 The location distribution of detector stations

IV. CASE STUDY

A signal intersection constructed by the Ronghua middle-road and the Rongjing road in Beijing is selected as the test bed. The intersection locates at the arterial section of Beijing economic-technological development area. The location and map are shown in Fig. 2. Field data were gathered at evening peak period (18:15-19:15) in July 25, 2011. The traffic volumes were video-taped by four cameras, which were located at the Station 5-Station 8 in Fig. 2, respectively. We recorded the maximum queue length at the entrance (i.e., Station 5-Station 8 in Fig. 2) into the intersection every signal cycle and use the mean of the above queue length as the hourly maximum queue length. The travel time was collected by floating cars. Fig. 2 shows the start and end points of travel time collection detectors, e.g., the Northwest-bound travel time collection is from Station 2 to Station 4, the Southeast-bound travel time collection is from Station 4 to Station 2.

The traffic model used is VISSIM Version 5.30. After building the VISSIM model, we apply the OGA to the parameter calibration. The default size of initial population D is 10. When the difference of maximum fitness values of two consecutive iterations is no more than 0.005 or the maximum number of iterations is 50, the program stops. Table 2 lists the relative errors of traffic volumes between VISSIM output results and field data at four entrances respectively, which shows that the values are all no more than 2.21%.

TABLE 2 VOLUME CALIBRATION OF THE INVESTIGATED INTERSECTION

	Simulation Results (vehicle/h)	**Field Data (vehicle/h)**	**Relative Error**
Northeast-bound	1173	1170	0.26%
Southwest-bound	486	497	2.21%
Southeast-bound	1766	1782	0.9%
Northwest-bound	525	528	0.57%

The proposed method is compared with calibration with the GA method and the orthogonal design method. In GA-based parameter calibration experiments, the fitness function, the size of initial population and the stopping rules are consistent with the OGA-based parameter calibration experiments. In the orthogonal design method, level number Q is 9, and the weight coefficient is 0.5. In the orthogonal design experiments, the parameter set is quantized by orthogonal design. The level number Q is the same with the OGA and the GA. The orthogonal array is constructed using Algorithm 1. Using the Q^J = 81 parameter groups as the VISSIM input data and selecting the maximum fitness value as the final output results.

We compare the maximum fitness values corresponding to four methods (i.e., default value, OGA, GA and the orthogonal method). The RMASE of travel time, RMASE of maximum queue length and fitness values are listed in Table 3. The maximum fitness value of OGA is 19.43, which is much bigger than that of other methods.

TABLE 3 COMPARATIVE ANALYSIS OF THE FITNESS FUNCTION VALUE

Methods	Parameter values	TT RMASE	MQL RMASE	Weight	Fitness
Default	[2,2,3,-4]	0.0607	0.3372	0.5	5.03
OGA	[1.125,0.5, 2.25,-2.75]	0.0225	0.0805	0.5	19.43
GA	[1.2,2.6, 1.33,-6]	0.0287	0.1380	0.5	11.99
Orthogonal method	[0.5,0.813, 1.625,-5]	0.0430	0.1378	0.5	11.06

Calculate the average errors for maximum queue length and travel time, respectively. The formula is as follows:

$$\varepsilon_i = 0.25\sum_{t=1}^{4}\frac{|d_{nt}^{s} - d_{nt}^{0}|}{d_{nt}^{0}} \quad n \in \{1,2\} \qquad (4)$$

where,

ε_1 = the average errors of TT

ε_2 = the average errors of MQL

d_{nt}^{s} =simulation output of *t*th entrance

d_{nt}^{0} = field data of *t*th entrance

t=1, 2, 3, 4 represent four entrance, i.e., northeast-bound, southwest-bound, southeast-bound and northwest-bound.

Fig. 3 and Fig. 4 represent the relative error of each access respectively. The relative error of the proposed method is less than that of other methods. Meanwhile, the relative error of travel time is higher than that of queue length, because investigators collect the queue length according to estimating the length of vehicles. It is obviously that the error must be existed.

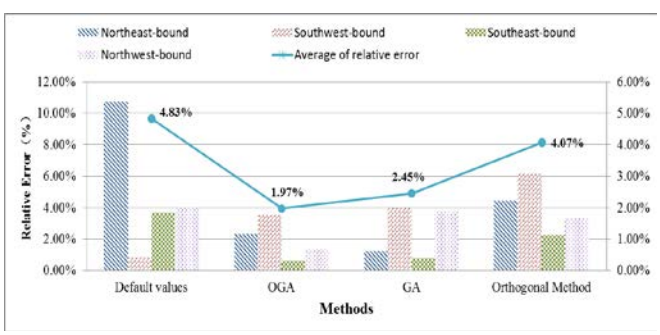


Figure 3 Comparison of TT relative error

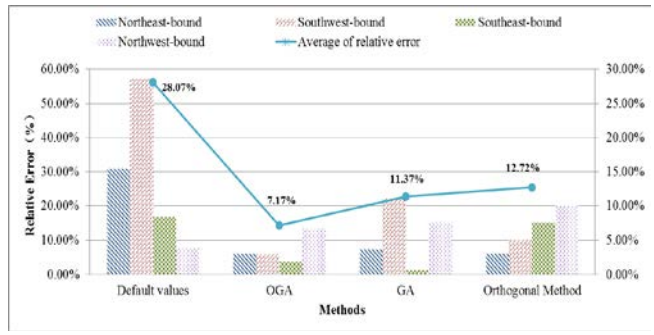


Figure 4 Comparison of MQL relative error

Fig.5 shows the comparison of convergence performance between the OGA-base calibration method and the GA-based calibration method. The former method stops when the iteration count is 10. Otherwise, the GA-based calibration method stops when iteration count is 40. To the OGA-base calibration method, it is obviously that constructing initial population consumes the most part of consuming time, because the program needs to run VISSIM 9^2 = 81 times to generate the output results. So we count the number of running VISSIM. Throughout the procedure, the proposed method runs VISSIM 238 times, and the later method runs 400 times.

Considering how the weight coefficient value impacts the calibration results. Fig.6 shows the profile of the fitness values versus weight coefficient values (i.e., weight coefficient value = 0, 0.2, 0.5, 0.8, and 1.0, respectively). The figure shows that a significant correlation exists between the fitness value and the weight coefficient values. The bigger the weight coefficient value is, the bigger the fitness value is. Considering the case of weight coefficient = 1.0, which means taking travel time into account indicator only, the fitness value is 62.70. Maximum queue length may reduce the fitness value, because field data collection exist big error.

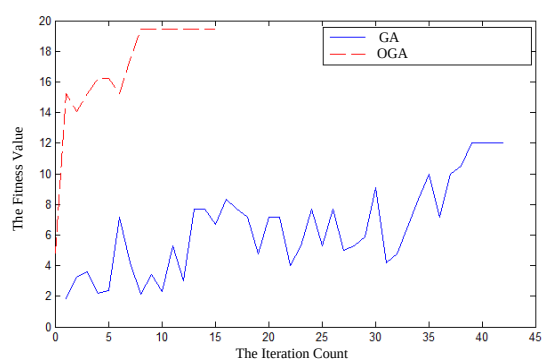


Figure 5 Comparison of convergence rate between two algorithms

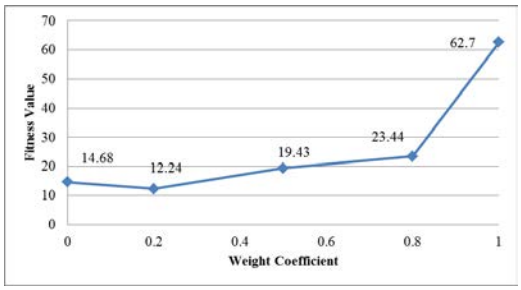


Figure 6 Comparison of the fitness value versus weight coefficient

V. CONCLUSION

This paper focuses on the automatic calibration method of traffic micro-simulation. A parameter calibration method based on orthogonal genetic algorithm is proposed. The first step, the process and the pseudo code of the OGA-based micro-simulation calibration method are given. The second step, we apply the proposed method to a signal intersection in Beijing. In the case study, the microscopic traffic flow simulation model VISSIM is selected. We compare the OGA method with the GA method and the orthogonal design method, respectively, i.e., the maximum fitness value, the relative errors of TT and MQL, the number of iteration. Experiment results show that the OGA outperforms the GA and the orthogonal design method in calibration. This paper also analyzes how the weight coefficient impacts the calibration results. A significant correlation exists between the fitness value and the coefficient values.

ACKNOWLEDGMENT

This research was supported by the National Science & Technology Support Program of China through Grant No. 2014BAG01B04.

REFERENCES

[1] D. Sun, L. Zhang, and F. Chen. "Comparative study on simulation performances of CORSIM and VISSIM for urban street network". Simulation Modeling Practice and Theory, Vol.37, pp. 18-29, 2013

[2] H. Filmon, and P. Luis Guilherme. "Sensitivity analysis of VISSIM driver behavior parameters on safety of simulated vehicle and their interaction with operations of simulated traffic". Transportation Research Board 92nd Annual Meeting, Washington DC, 2013

[3] F.C. Fang, and L. Elefteriadou. "Some guidenlines for selecting microsimulation models for interchange traffic operational analysis." Journal of Transportation Engineering, Vol. 131, No. 7, pp. 535-543, 2005

[4] I. I. Otković, T. Tollazzi, and M. Šraml. "Calibration of microsimulation traffic model using neural network approach. Expert Systems with Applications", Vol. 40, pp. 5965-5974, 2013

[5] B. Park, and J.D. Schneeberge. "Microscopic simulation modeling calibration and validation: a case study of VISSIM for a coordinated actual signal system". Transportation Research Record: Journal of the Transportation Research Board, pp. 185-192, 2003

[6] C. Gagnon, A.W. Sadek, and A. Touchette. "Calibration potential of common analytical and microsimulation roundabout models: new England case study". Transportation Research Record: Journal of the Transportation Research Board, pp. 77-86, 2008

[7] S. Kim, W. Kim, and L.R. Rilett. "Calibration of microsimulation models using nonparametric statistical techniques". Transportation Research Record: Journal of the Transportation Research Board, pp. 111-119, 2005

[8] J. Asamer, H. J.V. Zuylen, and B. Heilmann. "Calibrating car-following parameters for snowy road conditions in the microscopic traffic simulator VISSIM". IET Intelligent Transport System. Vol. 7, pp. 114-121, 2011

[9] L. Jie, H. V. Zuylen, Y. Chen, F. Viti, and I. Wilmink. "Calibration of a microscopic simulation model for emission calculation". Transportation Research Part C, Vol. 31, pp. 172-184, 2013

[10] B. Park, and H. Qi, "Development and evaluation of a procedure for the calibration of simulation models". Transportation Research Record: Journal of the Transportation Research Board, pp.208-217, 2005

[11] T. Ma, and B. Abdulhai. "Genetic algorithm-based optimization approach and genetic tool for calibrating traffic microscopic simulation parameters". Transportation Research Record 1806, pp. 6-15, 2002

[12] L.O. Kim, and L.R. Rilett. "A genetic algorithm based approach to traffic micro-simulation calibration using ITS data". Transportation Research Board 83rd Annual Meeting, Washington, D. C. 2004

[13] S. Menneni, C. Sun, and P. Vortisch. "Microsimulation calibration using speed-flow relationships". Transportation Research Record: Journal of the Transportation Research Board, pp.1-9, 2009

[14] R. L. Cheu, X. Jin, K.C. Ng, Y.L. Ng, and D. Srinivasan. "Calibration of FRESIM for Singapore expressway using genetic algorithm". Journal of Transportation Engineering, Vol. 124, pp. 526-535, 1998

[15] J. Ma, D. Hu, and H. Dong, and H. Zhang, "Calibration of microsimulation with heuristic optimization methods". Transportation Research Record: Journal of the Transportation Research Board, pp. 208-217, 2007

[16] K. Aghabayk, M. Sarvi, W. Young, and L. Kautzsch. "A novel methodology for evolutionary calibration of VISSIM by multi-threading". 36th Australasian Transport Research Forum, Australia, 2013

[17] Y. Wang, H. Liu, Z, Cai, and Y, Zhou. "An orthogonal design based constrain evolutionary optimization algorithm. Engineering Optimization". vol.39,pp.715-736, 2007

[18] Y. Leung, and Y. Wang. "An orthogonal genetic algorithm with quantization for global numerical optimization". IEEE Transaction On Evolutionary Computation , vol.5,pp.41-52, 2011

[19] Brackstone, M., and M. McDonald. "Car-following: a historical review". Transportation Research Part F, pp.181-196, 1999

[20] Manijunatha, P., P. Vortisch, P, and T. V. Mathew. "Methodology for the Calibration of VISSIM in Mixed Traffic", Transportation Research Board 92nd Annual Meeting, Washington D. C., 2013

Proceedings of

Distributed Multimedia Systems

A Novel Priority-based Deadlock Detection and Resolution Algorithm in Mobile Agent Systems

Wei Lu[1], Yong Yang[1,*], Liqiang Wang[2], Weiwei Xing[1], Xiaoping Che[1]
[1] School of Software Engineering
Beijing Jiaotong University
Beijing, China
[2] Department of Computer Science
University of Central Florida
Orlando, USA
Email: [1]{luwei,12112088,wwxing,xpche}@bjtu.edu.cn
[2]{lwang}@cs.ucf.edu

Abstract

Deadlock detection and resolution is one of the challenges in mobile agent systems, especially, when concurrent execution (i.e., more than one algorithm instances executing simultaneously) of algorithm instances. In this paper, we propose a deadlock detection and resolution algorithm in mobile agent systems. Priority-based approach is adopted in our algorithm to coordinate concurrent execution of algorithm instances. The liveness and safety properties of our algorithm are proved. Analysis and simulation results indicate that our algorithm can provide better performance and avoid duplicate detection and resolutions of the same deadlock in condition of concurrent execution.

Deadlock detection; Deadlock resolution; Mobile-Agent system; Distributed system; Concurrent coordination

1 Introduction

Rapidly expanding of available on-line information increases the need for bandwidth to support the industry infrastructure and bandwidth optimization [13]. For finding solutions of high bandwidth demand, many schemes are proposed in the past years.

Especially, Remote Procedure Calls (RPC) [7], Remote Programming (RP) and Code on Demand [3] are proposed to reduce network load in distributed systems before the emergence of mobile agent. Mobile agent with property of mobility greatly enhances the productivity of each computing element in the network and creates a uniquely powerful computing environment. A mobile agent acts as a self-contained software element responsible for executing a programmatic process, which is capable of autonomously migrating through a network.

There is no standard definition of mobile agent (some definitions can be found in literature [4, 10, 11, 13, 20]). Even so, we can obtain some common characteristics of mobile agent as summarized in [5]: mobility, autonomy, sociality, reactivity, proactivity, data acquisition, and route determination, etc.

Despite of many practical benefits, mobile agent technology also introducing new challenges (e.g., deadlock, rendezvous, and leader election). In addition, problem will become more complicated when concurrency error and considered [9, 14, 15]. In this paper, we propose a priority-based deadlock detection and resolution algorithm that can support concurrent execution of algorithm instances and provides better performance in mobile agent systems. Our algorithm can be roughly divided into two phases: decentralized information collection phase and centralized deadlock detection and resolution phase.

The paper is organized as follows: in Section 2, we review some related works. Description of our algorithm is presented in Section 3. Section 4 proves our algorithm. Section 5 illustrates the theoretical analysis and simulation results. Section 6 is the conclusions and future works.

2 Related Works

Some related works on deadlock detection and resolution in mobile agent systems as [1, 2, 6, 8, 16, 18, 19, 21].

Ashfield proposed a path pushing similar approach to detect deadlock in mobile agent system in [1, 2]. A consume

DOI reference number: 10.18293/DMS2016-014

agent creates a shadow agent to monitor its activities when it requests an exclusive lock on a resource. Shadow agent will create detection agent to perform deadlock detection if the consume agent has been block for a predetermined time interval. Detection agent responds for constructing WFG and detecting deadlock until a deadlock is detected by visiting related shadow agents.

In [8], Hosseini et. al, presented an improved algorithm based on Ashfield [1, 2]. They assigned a priority to both resource and consumer agent, and the detection agent created by consume agent with the least priority will suspend other detection agents that have larger priority. Therefore, only one detection agent (the lowest priority one) will pass through the whole cycle and can return.

Different with [1, 2], Elkady proposed an edge chasing similar algorithm to detect deadlock in mobile agent system in [6]. The proposed algorithm asserts a deadlock when a detection agent visiting a consume agent twice. In addition, the proposed algorithm supports deadlock avoidance and multiple detection agents.

Yang proposed two path pushing based deadlock detection algorithms named: MA-WFG and Host-WFS in [21]. In MA-WFG, mobile agent trying to construct a WFG to detect loop structure. Local construct WFG will pass to the mobile agent which locks the required resource to construct its local WFG. A deadlock is detected if there is loop topology in the collected WFG. To reduce the number of agent movements and relieve the network load, Host-WFS makes the WFG arranged in a form of "wait for set" that are distributed and stored on hosts. Therefore, the hosts control the termination of the path pushing, and mobile agent need not participate in the operation of path pushing.

Mani and Moshirpour, et.al, proposed an UML model based approaches to detect deadlock in Multi-Agent manufacturing systems in [16, 18]. They propose a behavioral model in form of UML 2.0 sequence diagrams from the modeling artifacts of the Multi-agent Software Engineering (MaSE) [17] to analysis and detect deadlock in multi-agent systems. Firstly, agents task diagrams are built during the analysis and designing of the MaSE methodology. The diagrams illustrate the activities performed by several cells (e.g., machines and robots). Each task diagram is a UML-like state chart diagram and will be converted to an UML activity diagram. Control Flow Paths (CFP) will be derived from the task diagrams. Then, resource requirement information can be gathered from each CFP that will be map into a machine requirement table. The table records the information and tasks of each agent with CFP and required machines (resources) of each CFP. During the run-time, each CFP that is running by an agent can initiate deadlock detection to query when blocked for a predefined time interval. The query will be sent to its direct depended CFP and propagated to indirect CFPs. The initiating CFP will determine itself deadlocked if it never receives a message from other CFPs to inform changing of the blocked status.

There are some limitations in previous works, such as: either no local deadlock detection or infinite deadlock detection, probability of false results or state explosion, performance inefficiency, high complexity, etc. To improve the performance efficiency and provide better solution in condition of concurrent execution. We propose a priority-based deadlock detection and resolution algorithm with better concurrent execution supporting in this paper.

3 Proposed Approach

3.1 Premises and Assumptions

Some properties in mobile agent systems as follows.

Network organization independence: neither the nodes nor the agents need to maintain knowledge about the size or topology of the network. The deadlock detection and resolution algorithm should not depend on a particular organization of the nodes in a mobile agent system.

Agent movement: agents must be allowed to lock resources and move for performing additional tasks.

Fault tolerance: messages or steps could be lost and gracefully recovery when it occurs.

Limit coupling: mobile agents that consume and manipulate resources in a mobile agent system should be separated from the deadlock detection and resolution process.

Based on the above properties of the mobile agent systems, we give some assumptions as:

- Agents can move through the network without lost, and network topology keeps static once the deadlock detection algorithm starts.
- A consume agent can choose to blocked when its resource lock request is rejected, or neglect the rejection and perform other tasks.
- There is no phantom deadlock based on the using of standard deadlock avoidance techniques, such as 2PC (two phase commit) or priority transactions.
- Host environment is the ultimate authority that can allow or deny the lock request of a resource.
- Host environment will be notified if a consume agent moves to another host environment.
- Agents and host environments are uniquely identified in the system through techniques such as static path proxy or naming service.
- Each consume agent can request at most one resource at a time, and will not request resource any more when blocked. It means the resource request model in this paper is single request model [12].

3.2 Algorithm Overview

We apply priority-based and edge chasing technique, that are commonly used in traditional deadlock detection algorithm in the distributed systems, in our algorithm.

Three kinds of mobile agents and the host environment are used in our algorithm descried as following. If a CA creates a DA or a RA, we call the CA hosting agent and the DA or RA spawned agent.

- **CA**, consume agent that can performs common tasks and exclusively locking needed resources. It spawns DA(s) and RA(s) to initiate deadlock detection and publish deadlock resolution. It does not have an active role in deadlock detection and resolution procedure. It reports its new location to the HE(s) that host resources locked by this CA when it move to a new HE.

- **DA**, detection agent that is spawned by CA and responsible for deadlock detection and resolution. It visits HEs to collect information of related CA(s) and construct global WFG. It responsible for detecting and resolving deadlocks according to the global WFG.

- **RA**, resolution agent spawned by CA used to notify the deadlock resolution to the victim CA.

- **HE**, host environment hosts mobile agents and provides APIs to hosted agents to interactive with itself. It controls and coordinates resource locking and unlocking. It provides information of hosted resources and mobile agents to DAs in deadlock detection activity.

3.2.1 Algorithm Initiation

A CA is assigned an unique id value when it is created in the mobile agent systems. It moves to a HE and requests resource to this HE when the needed resource is hosted by the HE. HE has the authority to grant or reject the request from CA. HE rejects the request if the resource is being locked by another CA and notifies the failure of resource requesting to the requesting CA. A CA can choose to blocked (waiting for the request resource to be unlocked) or discard this request to perform other tasks. The CA transfers from active to blocked state if it chooses blocked and wait.

The blocked CA initiates a deadlock detection and resolution algorithm instance by creating a DA after being blocked for an explicit time interval. The spawned DA has unique priority with the same value as the id of its hosting CA. Spawned DA starts distributed information collection phase by moving to the HE that hosts the needed resource and asking for the location of the CA that is locking this resource. Note that, it is possible that more than one DAs that roaming in the mobile agent system at the same time.

3.2.2 Deadlock Detection

The DA communicates with the HE to collects the resource locking information (i.e., wait-for relationship) of a CA, and determines whether continuing its visiting or not according to the situation of current visiting CA. A CA can not create a DA after been visited by other DA(s) until its request resource is granted.

A CA may have two kinds of situations when a DA arriving the HE that hosts this CA:

- Case1: the CA **has not** initiated a deadlock detection algorithm instance yet (e.g., no DA has been created by this CA when its HE is visited);

- Case2: the CA **has been** initiated a deadlock detection algorithm instance by creating a DA, and the detection process has been not terminated.

There are some possible sub-conditions in Case1 and Case2 of this CA:

- Case1-1: it is in an **active** state (e.g., this CA is not blocked by any resource at this moment);

- Case1-2: it is in a **blocked** state, but has been not initiated a deadlock detection instance by creating DA.

- Case2-1: its spawned DA has a **larger** priority than the priority of visiting DA.

- Case2-2: its spawned DA has a **lower** priority than the priority of visiting DA.

- Case2-3: its spawned DA has a **same** priority with the priority of visiting DA.

In Case1-1, the visited CA is in active state. In this condition, the visiting DA will determine that there is no deadlock and return to its hosting CA with information of visited CA(s). The hosting CA will make the decision of no deadlock and wait for locked resource to be available.

In Case1-2, different with Case1-1, the visited CA is blocked for resource that is locked by another CA. The visiting DA will move to the next HE. This procedure will continues till one of the following conditions is meet. One is that a current visited CA is in a active state like Case1-1, another is that this DA visits a CA that has initiated an algorithm instance by spawning a DA with a different priority, this falls into the cases of Case2.

In Case2-1, the visiting DA visits a CA that has been initiated an algorithm instance by spawning a DA with a higher priority. In this case, the visiting DA adds information of the visited CA into its visiting list and return to its hosting CA. The hosting CA will determine whether performing a second round of detection or not according to the visiting list collected by its returning DA.

In Case2-2, the visiting DA will not return until the visited CA receives the returning of its own spawned DA. The visiting DA combines its own visiting list and the visiting list of the visited CA together, and then, returns to its hosting CA with the combined visiting list.

In Case2-3, this condition represents the existence of deadlock when a visiting loop forms.

In Case1-1, Case1-2, and Case2-3, an explicit decision of deadlock can be made. However, there needs some other steps to continues the deadlock detection in Case2-1 and Case2-2 after returning of spawned DA(s) with visiting list.

After the returning of spawned DA(s) in Case2-1 and Case2-2, the hosting CA(s) will locally record the visiting list and make determination of a second round detection. A CA will launch the next round of detection if the following two conditions are satisfied:

- Condition-1: this CA has the largest id among the CAs in the local visiting list;
- Condition-2: this CA has been visited by other DAs.

The spawned DA will execute a next round of visiting with a same priority as the last round. In addition, the target visiting CA is the rear CA of the collected visiting list. The rest procedure is the same as Case2-1 or Case2-2 till a explicit decision of deadlock can be make, then, the detection process terminates.

3.2.3 Deadlock Resolution

A global WFG can be constructed after the first phase terminating. The second phase starts detecting and resolving deadlock if there exists a deadlock. Proposed algorithm can guarantee that only the DA with the highest priority will detect the deadlock at last. This DA will select a victim CA to release its locked resource to break the deadlock (simply, itself). The hosting CA which spawn a RA and send this RA to the victim CA for deadlock resolution. The victim CA releases the locked resource after receiving the RA. Then, the deadlock is resolved, and all blocked CA(s) can perform the next tasks.

3.3 Formal Description

Formal description of our algorithm is given in this section. Variables used in our algorithm are given in Table 1.

Algorithm 1 illustrates that how to handle the arriving of a detection agent in a visited CA (the possible cases in the previous subsection). Algorithm 2 gives the description of the strategy when a DA returning to its hosting CA.

3.4 Case Study

As shown in Figure 1,there are 6 HEs ($HE1 - HE6$), 6 CAs ($CA1 - CA6$), and 15 resources ($R1 - R15$) in a

Table 1: Variables used in the proposed algorithm.

Variable	Descriptions
CA_i	Consume agent with a unique $id = i$, the hosting CA of its spawned DA_i.
DA_i	Detection agent created by CA_i with a priority i, spawned DA of CA_i.
DA_i^{VL}	Visiting list recorded in DA_i. It is ordered in visited order. Such as,$\{CA_1, CA_2, ..., CA_k\}$ means the visiting order is $CA_1 \rightarrow CA_2 \rightarrow, ..., \rightarrow CA_k$.
VL_i	Visiting list recorded at CA_i.
RA_i	Resolution agent created by a consume agent with $id = i$. It is used to notify the victim consume agent of deadlock resolution.

Algorithm 1 when DA_i arriving at CA_j

$VL_j = DA_i^{VL} \cup VL_j$;
$DA_i^{VL} = DA_i^{VL} \cup VL_j$;
if CA_j has initiated a deadlock detection instance **then /*Case 2*/**
  if $DA_j < DA_i$ **then /* *Case 2-1*/***
    waiting till the returning of DA_j;
    $DA_i^{VL} = DA_i^{VL} \cup VL_j \cup DA_j^{VL}$;
    DA_i return to CA_i with DA_i^{VL};
  else if $DA_j > DA_i$ **then /* *Case 2-2*/***
    $DA_i^{VL} = DA_i^{VL} \cup VL_j$;
    DA_i return to CA_i with DA_i^{VL};
  else/* *Case 2-3*/*
    a deadlock is detected;
    creating a RA_i and sending RA_i to the CA that is selected as victim by DA_i to resolve the deadlock;
  end if
else/*Case 1*/
  if CA_j is active **then /* *Case 1-1*/***
    $DA_i^{VL} = DA_i^{VL} \cup \{CA_j \rightarrow \varnothing\}$;
    DA_i return to CA_i with DA_i^{VL};
  else/* *Case 1-2*/*
    $DA_i^{VL} = DA_i^{VL} \cup \{CA_j \rightarrow CA_k\}$;
    DA_i moves to the next CA_k;
  end if
end if

Algorithm 2 when DA_i returning to CA_i with DA_i^{VL}

$VL_i = VL_i \cup DA_i^{VL}$;
if there is cycle in VL_i and DA_i has the largest priority among all priority values of collected CAs **then**
  a deadlock is detected;
  creating a RA_i and sending RA_i to the victim CA;
else
  if DA_i is the largest in VL_i and HE of CA_i has been visited by other DA **then**
    DA_i moves to CA_j;
  else
    while the requested resource is not granted **do**
      waiting the granting of resource;
      if receiving a RA_i **then**
        releasing the locked resource, and break the while loop;
      end if
    end while
  end if
end if

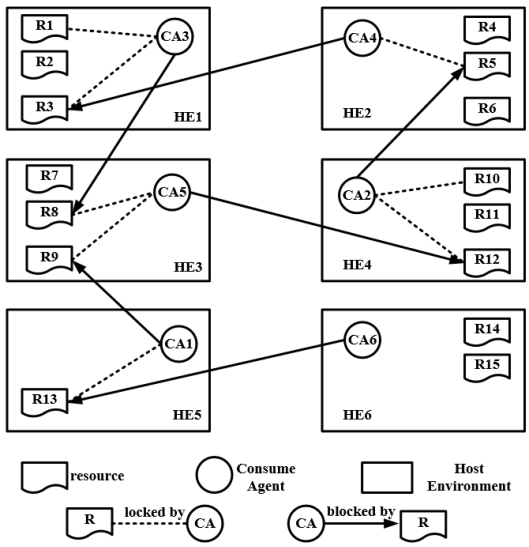


Figure 1: Instantaneous state of a mobile agent system.

mobile agent system. 8 of the 15 resources are locked. Assuming that Figure 1 is a transient state of the mobile agent system. Take HE1 in Figure 1 for example, HE1 holds three resources ($R1$, $R2$,and $R3$) and a Consume Agent ($CA3$). $CA3$ has locked resource $R1$ and $R3$, and is blocked by $R8$ which is locked by $CA5$ locates at $HE3$.

Assuming that Consume Agents $CA1$, $CA2$, $CA3$, $CA4$, $CA5$, and $CA6$ have been blocked for a specific time interval. Therefore, they will spawn DAs to execute deadlock detection and resolution. We also assume that the time of a DA's movement is finite but uncertain, and the DAs will be not lost or tampered.

One possible scenario is all consume agents creating detection agent and initiating deadlock detection simultaneously. In this condition, detection agents $DA1$, $DA2$, $DA3$, $DA4$, $DA5$, and $DA6$ will visit $HE5$, $HE4$, $HE1$, $HE2$, $HE3$, and $HE6$ to collect information.

$DA1$ will return to $HE1$ immediately with a new resource request (i.e., wait-for) information of $CA5$ ($\{CA5 \rightarrow CA2\}$) when $DA1$ arrives at $HE3$ before returning of $DA5$. Because $DA1$ has a lower priority than $DA5$ that is spawned by $CA5$. $DA2$ and $DA3$ returns with ($\{CA4 \rightarrow CA3\}$) and ($\{CA5 \rightarrow CA2\}$). Assuming that $DA4$ arrives at $HE1$ and returns to $CA4$ before the returning of $DA3$, therefor, $DA4$ will collect the information of $CA3$ as ($\{CA3 \rightarrow CA5\}$) and $DA6$ return after the returning of $DA2$ and $DA1$, the visit lists of $DA5$ and $DA6$ will be ($\{CA2 \rightarrow CA4 \rightarrow CA3\}$) and ($\{CA1 \rightarrow CA5 \rightarrow CA2\}$), respectively.

The local visiting list in each consume agent are illustrated in Table 2 after the above steps.

Only $DA5$ will launch a second round of detection, because the id of it is the largest among CAs in the local visiting list. $DA5$ will move to $HE3$ for collecting wait-for information of $CA3$. The other consume agents will suspend to wait for the detection and resolution results. As the same rule in the first round, the local visiting list will be changed as shown in Table 3 after the second round returning of detection agents

Table 2: Local visiting list in CAs after returning of their spawned DAs.

VL_i	**Local visiting list**
VL_1	$\{CA1 \rightarrow CA5 \rightarrow CA2\}$
VL_2	$\{CA2 \rightarrow CA4 \rightarrow CA3\}$
VL_3	$\{CA3 \rightarrow CA5 \rightarrow CA2\}$
VL_4	$\{CA4 \rightarrow CA3 \rightarrow CA5\}$
VL_5	$\{CA5 \rightarrow CA2 \rightarrow CA4 \rightarrow CA3\}$
VL_6	$\{CA6 \rightarrow CA1 \rightarrow CA5 \rightarrow CA2\}$

Table 3: Local visiting list in CAs after returning of their spawned DAs.

VL_i	**Local visiting list**
VL_1	$\{CA1 \rightarrow CA5 \rightarrow CA2\}$
VL_2	$\{CA2 \rightarrow CA4 \rightarrow CA3\}$
VL_3	$\{CA3 \rightarrow CA5 \rightarrow CA2\}$
VL_4	$\{CA4 \rightarrow CA3 \rightarrow CA5\}$
VL_5	$\{CA5 \rightarrow CA2 \rightarrow CA4 \rightarrow CA3 \rightarrow CA5\}$
VL_6	$\{CA6 \rightarrow CA1 \rightarrow CA5 \rightarrow CA2\}$

A deadlock is detected by $DA5$ after its returning to $CA5$ according to the local visiting list. $DA5$ will select a victim CA that is involved the deadlock from the local visiting list. Then, deadlock resolution task will be conducted by $CA5$ by spawning a $RA5$ and sending it to the victim CA(e.g., $CA2$). $CA2$ will release the lock on resource $R12$ after receiving of $RA5$, and the deadlock will be broken.

4 Correctness Proofs

In this section, we prove the liveness and safety property of our algorithm. In addition, some other necessary theorems are also proved.

Theorem 1 *At any given time, there is at most one deadlock in a WFG.*

Proof 1 *Since the resource request model is single request model, the number of outgoing arc at each node is one at most. Assuming that there are more than one deadlock in the collection WFG. It means at least two cycles in the WFG and they are either independent or intersection with each other. Due to the single request model, it is impossible that two cycles are intersection. So, they should be detected in separated WFG rather than in one WFG. Therefore, there is at most one deadlock in one WFG at any given time.*

Table 4: Comparison of the worst performance in aspect of concurrent execution.

	[2]	[8]	[21]^1	[21]^2	[6]	Our
number of DA moves in detection	∞	$\frac{n(1+n)}{2}$	$\geqslant \frac{n(1+n)}{2}$	$\geqslant \frac{n(1+n)}{2}$	n^2	$4(n-1)$
number of RA moves in resolution	$-$	1	$-$	$-$	$\geqslant n^2$	1
number of transmitted CA information	∞	$\frac{n^2(1+n)}{2}$	$\geqslant \frac{n^2(1+n)}{2}$	$\geqslant \frac{n^2(1+n)}{2}$	$\frac{n^2(1+n)}{2}$	$\frac{n(1+n)}{2}$
Time delay	∞	n	n	n	n	$n+1$

[21]^1 and [21]^2 are the two algorithm proposed in literature [21]; $-$: no deadlock resolution or difficult to analysis theoretically; n: number of detection agent in a WFG.

Theorem 2 *If a deadlock exists, only one DA will detect it and one CA will resolve it.*

Proof 2 *According to Algorithm 1, only DA with the highest priority will collected the whole WFG and perform final deadlock detection. Therefore, only the hosting with the largest value of id will resolve a detected deadlock. The theorem is proved.*

Theorem 3 *No phantom deadlock will be detected and resolved.*

Proof 3 *The phantom deadlock appears when two deadlock detection and resolution algorithm instance attempt to perform on the same CA. By Theorem 2, there can be a maximum of one algorithm instance will detect and resolve a deadlock. Hence, there is no possibility for a phantom deadlock in our algorithm.*

Theorem 4 *(Liveness) If there is a deadlock in the system, it should be detected and resolved in finite time.*

Proof 4 *Let's assume that there is a deadlock in the system and not be detected when at least one of the CA initiate deadlock detection procedure. There are two possible situations, one is that the DA lost during its travel. This is not possible duo to the assumption that DA will not lost in the system. Another one is no deadlock detects when a DA with the highest priority collected WFG. This contradicts the assumption. Therefore, if the movement of agent is in finite time, the deadlock will be detected and resolved in finite time.*

Theorem 5 *(Safety) If a algorithm instance resolves a deadlock in the system, it is in fact in the system and resolved by only one algorithm instance.*

Proof 5 *By Theorem 3, we can prove that the detected deadlock actually in the system. By Theorem 1 and Theorem 2, we can derive that the detected deadlock will be detected and resolved by only one algorithm instance. The theorem is proved.*

5 Performance Analysis

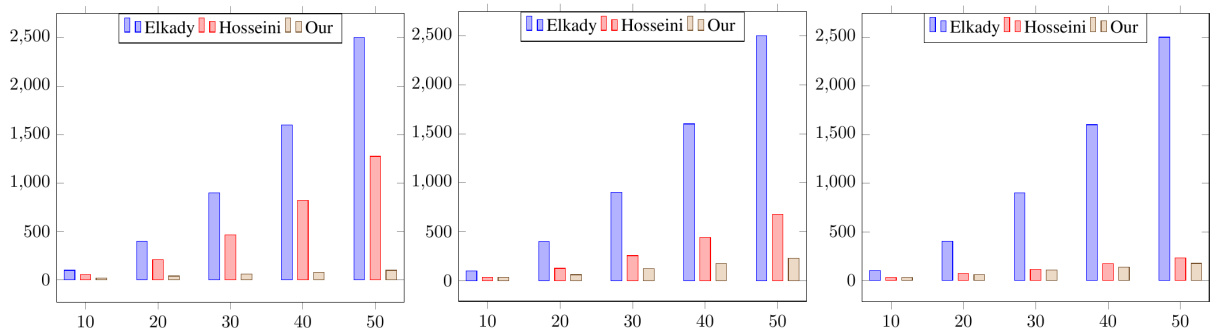
In this section, we analysis the worst case of performance among previous and our algorithms, and perform simulation each algorithm. It should be noted, only the situation of concurrent execution is considered in the analysis. This analysis contains number of DA moves in detection, number of RA moves in resolution, total number of transmitted CA information and time delay. Specifically, one piece of transmitted CA information means information of one CA (e.g., $CA_i \rightarrow CA_j$ as one piece information of CA_i).

5.1 Theoretical Analysis

In this theoretical analysis, we assume that n consume agents (CA_1, CA_2,...,CA_n) involved in a deadlock, and all consume agents initiate deadlock detection and resolution algorithm instances with arbitrary time order. Our algorithm has the worst case of DA movements in Table 5.B and the worst case of transmitted CA information in Table 5.A. In Table 5.B, the first round of deadlock detection will has $2n$ DA moves. The second round has $2 * (\frac{n}{2})$ DA moves and the i_{th} round has $2 * \frac{n}{2^{(i_{th}-1)}}$ DA moves. This procedure continues till $1 \leqslant \frac{n}{2^{(i_{th}-1)}} < 2$. Therefore, number of DA moves in detection is $2 * \frac{2*(1-2^{log_2^n})}{1-2} = 4(n-1)$. In Table 5.A, each DA_i will return to CA_i with $i-1$ pieces of CA (these CAs has lower id than i) information. Therefore, total number of transmitted CA information is $\frac{n(1+n)}{2}$.

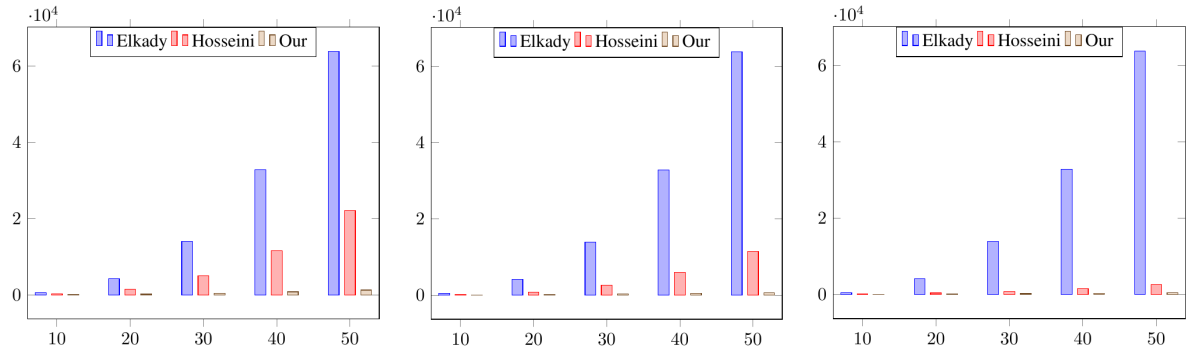
The DA with the highest priority performs deadlock detection after constructing WFG, and it will select a victim from the deadlocked CAs. Then, its hosting CA will create a resolution agent (RA) and send it to the victim consume agent. Hence, the resolution transmits only one time of RA to resolve the deadlock. If the priority of each DA in the cycle within a decreasing order and the least priority waits for the highest one, the worst case of time delay in our algorithm is n.

Based on the same criteria, Table 4 illustrates performance analysis results (in the worst case) between previous works and our work.



(a) Number of DA movements in Table 5.A. (b) Number of DA movements in Table 5.B. (c) Number of DA movements in Table 5.C.

Figure 2: Performance comparison on detection agent movement between some previous works and our work.



(a) Totle number of transmitted CA information transmitted in Table 5.A. (b) Totle number of transmitted CA information transmitted in Table 5.B. (c) Totle number of transmitted CA information transmitted in Table 5.C.

Figure 3: Performance comparison on information transmission between some previous works and our work.

5.2 Simulation Results

The simulation program is written in Scala (version 2.11.7) with Akka (version 2.3.11). We compares two previous works (Elkady and Hosseini [6, 8]) with our work, and run the simulation on three types of scenes as shown in Table 5. The program of each algorithm runs 100 times, and the result is the average value.

Table 5: Description of possible scenes in the simulations.

Type	Description
A	Each agent requests a resource that is locked by another agent, wait-for relationship as: $n \to n-1 \to ... \to 1 \to n$;
B	Each agent requests a resource that is locked by another agent, wait-for relationship as: $n \to 1 \to n-1 \to 2... \to \lfloor \frac{n}{2} \rfloor \to \lfloor \frac{n}{2} \rfloor - 1$;
C	Each agent request a resource that is locked by another agent, and each agent can be waited by one another agent.

We just statistic the number of DA moves and total number of transmitted CA information in detection phase.

Figure 2 and Figure 3 give the simulation results of performance comparisons. Figure 2 and Figure 3 illustrate the results of the comparisons in aspect of DA moves and transmitted CA information, respectively. Label on horizontal axis and ordinate axis represent the number of consume agents and results, respectively.

From the simulation results, we can find that the number of detection agent movements increasing linearly with the growth of the number of consume agent. The other two algorithms have an exponential growth. The total number of transmitted CA information is the same. The main reason is that our algorithm reuses the collected wait-for information of consume agents with lower id. This dramatically reducing the number of detection movements and total number of transmitted CA information during the detection procedure. In addition, another benefit of our algorithm is that only one agent will detect and resolve the deadlock. This can avoid all associated problems that is caused by multiple resolution of the same deadlock.

Note that, results of Hosseini has better performance than Elkady in Figure 2 and Figure 3. Actually, they

have the same theoretically performance complexity (see Table 4). However, the worst case of Hosseini's approach need extreme conditions of algorithm initiation time that is difficult to implement in this simulation. However, it may happen in real system.

6 Conclusions and Future Works

In this paper, we propose a priority-based deadlock detection and resolution algorithm in mobile agent systems. Our algorithm adopts priority-based approach to coordinate the concurrent execution of algorithm instances. Our algorithm guarantees that a deadlock can only be detected and resolved by one agent. This simplifies the resolution of deadlock, and no phantom deadlock will be detected. In addition, our algorithm provides better performance when concurrent executing of algorithm instances. We prove two important properties (liveness and safety) of our algorithm. Performance analysis and simulation results demonstrate that our algorithm has better performance in aspect of agent movements and information volume.

In the near future, we will pay more attention on the other complexity resource request models in deadlock detection and resolution of mobile agent systems. Besides that, fault tolerance of agent and host environment crashing will be another topic of research in our future works.

ACKNOWLEDGMENT

This work was supported in part by National Natural Science Foundation of China (No.61100143, No.61272353, No.61370128, No.61428201, No.61502028), Program for New Century Excellent Talents in University (NCET-13-0659), Beijing Higher Education Young Elite Teacher Project (YETP0583).

References

[1] B. Ashfield. *Distributed Deadlock Detection in Mobile Agent Systems*. PhD thesis, Carleton University Ottawa, 2000.

[2] B. Ashfield, D. Deugo, F. Oppacher, and T. White. *Distributed deadlock detection in mobile agent systems*. Springer, 2002.

[3] P. G. P. Carzaniga, A. and G. Vigna. Is code still moving around? looking back at a decade of code mobility. In *Companion to the proceedings of the 29th International Conference on Software Engineering*, pages 9–20. IEEE, 2007.

[4] M. M. K. G. Challenger, M. and T. Kosar. Declarative specifications for the development of multi-agent systems. *Computer Standards and Interfaces*, 43:91–115, 2016.

[5] V. Della Mea. Agents acting and moving in healthcare scenario - a paradigm for telemedical collaboration. *IEEE Transactions on Information Technology in Biomedicine*, 5(1):10–13, 2001.

[6] A. Elkady. *Mobile Agents Deadlock Detection in Absence of Priorities*. PhD thesis, Carleton University Ottawa, 2006.

[7] D. R. G. Fredriksson, Olle and B. Wheen. Towards native higher-order remote procedure calls. In *Proceedings of the 26nd 2014 International Symposium on Implementation and Application of Functional Languages*. ACM, 2014.

[8] R. Hosseini and A. T. Haghighat. An improved algorithm for deadlock detection and resolution in mobile agent systems. In *International Conference on Computational Intelligence for Modelling, Control and Automation, and International Conference on Intelligent Agents, Web Technologies and Internet Commerce*, volume 2, pages 1037–1042. IEEE, 2005.

[9] H. Huang, L. Wang, B. C. Tak, L. Wang, and C. Tang. Cap3: A cloud auto-provisioning framework for parallel processing using on-demand and spot instances. In *Proceedings of the 2013 IEEE Sixth International Conference on Cloud Computing*, CLOUD '13, pages 228–235. IEEE, 2013.

[10] D. Isern and A. Moreno. A systematic literature review of agents applied in healthcare. *Journal of medical systems*, 40(2):1–14, 2016.

[11] R. Jain, F. Anjum, and A. Umar. A comparison of mobile agent and client-server paradigms for information retrieval tasks in virtual enterprises. In *Research Challenges, 2000. Proceedings. Academia/Industry Working Conference on*, pages 209–213. IEEE, 2000.

[12] E. Knapp. Deadlock detection in distributed databases. *ACM Computing Surveys (CSUR)*, 19(4):303–328, 1987.

[13] D. Kotz and R. S. Gray. Mobile agents and the future of the internet. *Operating systems review*, 33(3):7–13, 1999.

[14] H. Ma, S. R. Diersen, L. Wang, C. Liao, D. Quinlan, and Z. Yang. Symbolic analysis of concurrency errors in openmp programs. In *the 42th International Conference on Parallel Processing*, pages 510–516. IEEE, 2013.

[15] H. Ma, L. Wang, and K. Krishnamoorthy. Detecting thread-safety violations in hybrid openmp/mpi programs. In *Proceedings of the 2015 IEEE International Conference on Cluster Computing*, pages 460–463. IEEE, 2015.

[16] N. Mani, V. Garousi, and B. H. Far. Search-based testing of multi-agent manufacturing systems for deadlocks based on models. *International Journal on Artificial Intelligence Tools*, 19(04):417–437, 2010.

[17] L. Moreau, M. Luck, S. Miles, J. Papay, K. Decker, and T. Payne. Methodologies and software engineering for agent systems. 2004.

[18] M. Moshirpour, N. Mani, A. Eberlein, and B. Far. Model based approach to detect emergent behavior in multi-agent systems. In *Proceedings of international conference on Autonomous agents and multi-agent systems*, pages 1285–1286, 2013.

[19] N. Sofy and D. Sarne. Effective deadlock resolution with self-interested partially-rational agents. *Annals of Mathematics and Artificial Intelligence*, 72(3–4):225–266, 2014.

[20] J. Waldo. Mobile code, distributed computing, and agents. *IEEE Intelligent Systems*, (2):10–12, 2001.

[21] J. Yang. *Design of fault tolerant mobile agent systems*. PhD thesis, Hong Kong Polytechnic University, 2006.

An Entropy based Product Ranking Algorithm using Reviews and Q&A Data

Bushra Anjum
Amazon Inc., 1194 Pacific St., San Luis Obispo
CA 93401, USA
banjum@amazon.com

Chaman Lal Sabharwal
Missouri University of Science and Technology, Rolla
MO 63128, USA
chaman@mst.edu

***Abstract* — Amazon.com, along with several other commercial websites for products and services, provides a platform for consumers to share their opinions by providing reviews and answering product related questions (QA data). These opinions can be quantitative, qualitative or a combination of both. Owing to the large corpus of such data available, there are several learning and classification approaches available to scrutinize them e.g., those based on Entropy measures, machine learning, stochastic, and natural language processing etc. In this paper, we review some of the prominent techniques and explore a hybrid approach, involving Entropy, Bilinear and statistical measures, to use heterogeneous consumer data and simultaneously analyze and rank products for customers. With experimental results, we show that our approach effectively ranks products using (1) text reviews (2) QA data and (3) star rating of products. We also make a case that the ranks calculated are more relevant to the customers and can enable better prediction on the products sale for the sellers.**

Keywords-product reviews, product ranking, similarity, classification

I. INTRODUCTION

Gaining insights from product reviews has emerged as a novel field of research and has valuable implications in the real world. Many e-commerce websites, such as Amazon.com, provide a platform for consumers to share their opinions. Unbiased reviews by other consumers build confidence of a customer to go ahead with a transaction [1]. The reviews are generally quantitative in the form of star rating, or qualitative in the form of comments written in plain text. Due to the large corpus of review data available and limited customer time, researchers are not only focusing on means to ascertain the quality, authenticity, and usefulness of reviews but also on ranking products based on the available data.

It has been ascertained that the product rating with stars (or numerical scale 0-5 etc.) alone does not provide enough semantics information about customer's sentiment and that a text based review is more revealing in that context [2]. Reviews are subjective opinions and judgment about a product or the service. Hence, nowadays the trend is to use both star based ranking and text reviews in all types of surveys pertaining to products and services. However, it is also important to note that though it is quick and easy to process quantitative ratings, producing qualitative semantic information is a challenging problem because the deciphering and evaluation of text reviews is both time-consuming and technically complex due to an unstructured form of natural language text [1].

As mentioned earlier, it is important and necessary to classify the relevance of product related information, for the sake of consumer's premium time [3]. Reviews are a major source of such information. Another interesting area of research, though not a focus of this work, is the availability of search and evaluation tools that helps a customer distinguish between forged and genuine reviews from their implied reputation [4]. In this work, we are assuming that the reviews provided are genuine with varying degree of relevance

There are several approaches used to ascertain the quality, authenticity, and usefulness of the reviews such as Entropy (H), Decision trees (DT), Singular Value Decomposition (SVD), support vector machines (SVM), machine learning (ML), stochastic probability, and natural language processing (NLP) etc. These approaches are explained further in the related work section and help in evaluating the helpfulness of a review, rank products based on relevance from the reviews, enable prediction on the products, and reduce the search time.

DOI reference number: 10.18293/DMS2016-024

In this paper, we will explore a hybrid analysis approach to use heterogeneous product review data (star ranking, text based reviews, question/answer data) to simultaneously analyze and evaluate the reviews. We show that adapting the ranking algorithm by simultaneously using Entropy and Bilinear Similarity measures (explained later in the paper) yields more accurate evaluation than using them in isolation and on restricted data types. With experimental results, we learn that our approach is effective in ranking products based on interest, and relevance and also accurately relies on review text and question/answer data to produce the final ranking of the product.

The paper is organized as follows: Section II describes the preliminary background and literature review, section III describes our contributions in detail, section IV is on experiments and section V lists conclusions followed by references.

II. Background and Related Work

A. Preliminaries

Here we describe all the *standard* terms used in this wok and discussion. Data instance/object is in the form of a vector. We assume all vectors are *column* vectors. A matrix represents the aggregate of all data instances. A *row* vector is the *transpose* of a column vector. A *unit* vector is a vector of length one. A *weight* vector is a vector with positive components whose sum is *unity*. The *dot* product of two vectors is a scalar, *Hadamard* product of two vectors is a vector of point-wise products of the corresponding components of the two vectors. The *similarity* between two *unit* vectors can be defined by *simple* dot product, *weighted* dot product, and *complex* measure Okapi *BM25 (Best Matching 25)* [5], see section II.C. The data matrix format is also called *TF-IDF* format, meaning term-frequency by inverse-document-frequency. *Entropy* is the measure of uncertainty in the classification, where the smaller the value of uncertainty, the better is the resulting classification. Accuracy is measured as: *Precision*, how accurate prediction is over only positive instances, *Recall*, how accurate positive prediction is for only relevant instances, and *F-measure*, which is the *weighted Harmonic mean* of Precision and Recall.

B. Related work

The field of machine learning is vast, challenging and, is increasingly being utilized by non-technical consumers, many of which do not have a basic understanding of the foundational principles. In our case, the task is to make a recommendation based on product reviews. At a high level, there are two kinds of tasks frequently seen in machine learning: *classification*, and *ranking*. There are several tools used to accomplish it, entropy and regression being two of them. For instance, a student in the class gets "A" grade, the letter grade is *classification*, and the numeric grade is determined by *an algorithm* for predicting the letter grade from multiple scores of tests, home works, and quizzes. The relative *ranking* of students in the same grade is determined from scores where the raw score lies in the scores range.

Also, it is a de facto reality that more complex machine learning metric is not necessarily an improved evaluation technology; especially when data is limited. Simple techniques frequently outperform more complex ones [3]. In order to evaluate the usefulness of reviews of a product, we want to build upon the available resources.

Several researchers have been working on sentiment classification using online product reviews with the goal to determine whether the consumers find the product useful. The machine learning for classification can be performed in two ways: supervised learning [6] to predict the semantics of adjectives, unsupervised learning [2] to classify reviews as helpful or no helpful by analyzing the semantic orientation of reviews. We use supervised machine learning here because there is a classification attribute in the data. The quality and helpfulness of the reviews is an active area of research.

For supervised learning, [7] developed a method using support vector machine (SVM) to automate the review helpfulness evaluation, [8] used entropy-based approach to explore the online review helpfulness and [5] used bilinear approach for classification of Amazon data.

The related area of research is the *effect* of online product reviews on the *product sales*. The consumers consider not only the reviews but also the reputation of the reviewer [9]. Finding a best data-mining tool is itself a data mining problem with no resolution.

C. Helpfulness Metrics for Supervised Learning

First, we briefly discuss the Entropy measure [8] and bilinear similarity [5] measures for relevance ranking of reviews for products. We will be using these measures in a hybrid fashion later in this paper. Our approach is unique as it uses the heuristics from singular value decomposition (SVD) and best of bilinear and entropy. The resulting metric may be used to determine relevance, classification and ranking simultaneously.

An online review consists of words, including opinion words, product features, product parts, and other words. The importance of each word to relevance and helpfulness of a review is calculated from the training dataset that has the consumer voting data information.

1) Helpfulness Model

Let P be the set of products, $p \in P$ be a product; R be the set of reviews and $r \in R$ be a review. Let r_h be the number of customers who vote for the review r as helpful. Let $r_{\bar{h}}$ be the number of customers who vote against the review r i.e. the review r is "not helpful". The review's helpfulness measure is defined by

$$H = \frac{r_h}{r_h + r_{\bar{h}}} .$$

Let τ = 0.66, say, be the threshold to indicate that the *review is helpful* if $H > \tau$, i.e. two thirds of the consumers liked it. This is a simplistic view of the metric for classifying helpfulness of online reviews. Two of the metrics, we use for our algorithm are information Entropy [8] and Bilinear similarity [5], both help in the overall ranking of the products.

2) Entropy Model

First, we derive formulation for general *n* categories and single item classification. Then we adapt it to a review consisting of several words for two categories: "helpful" and "not helpful". Let $C = \{c_1, c_2, c_3, \ldots, c_n\}$ be the set of categories for products in the review space. Then information Entropy [10] needed to classify a review *r* is

$$H(C) = -\sum_{i=1}^{n} P_r(c_i) \log P_r(c_i)$$

H stands in honor of Boltzmann's H-Theorem. Non-uniform entropy is normalized by the maximum entropy -$\log(\frac{1}{n})$ = log(n), the normalized entropy is

$$H(C) = -\sum_{i=1}^{n} \frac{P_r(c_i) \log P_r(c_i)}{\log n}$$

The amount of information contributed by a term *t* (or word relevant to a document) is

$$H(C|t) = -\sum_{i=1}^{n} \frac{P_r(c_i|t) \log P_r(c_i|t)}{\log n}$$

Information gain (entropy reduction by knowledge of *t*) is then

$$G(t) = H(C) - H(C|t)$$

Higher the gain, higher the ability to classify, lower the uncertainty in the ability to classify.

This is the standard metric used to classify unknown (new items, not in the training set) related items. To get a better estimate for prediction, we include the occurrence and non-occurrence of t

$$G(t) = H(C) - P_r(t)\, H(C|t) - P_r(\bar{t})\, H(C|\bar{t})$$

contributes to the prediction ability. Larger the value of G(t), better predictor by knowledge of *t.*

Note. In this case, "helpful" and "not helpful" are two categories. Let c_1 be a category of "helpful" and c_2 be "not helpful". In order to differentiate prediction ability for two categories, G(t) upgraded to

$$\text{Gain}(t) = \begin{Bmatrix} G(t) \; if \; P(c_1|t) > P(c_2|t) \\ -G(t) \quad otherwise \end{Bmatrix}$$

3) Prediction Computation

If there are *n* words in the review, then score is for a review is sum of the gain of each term

$$Score(r_i) = \sum_{k=1}^{N} Gain(t_k) * f(r_i, t_k)$$

where Gain(t_k) is k-th word's gain value, *N* is the number of words in review r_i and f(r_i, t_k) ,the Heaviside function,

$$f(r_i, t_k) = \begin{cases} 1 \; if \; term \; t_k occurs \; in \; r_i \\ 0 \; otherwise \end{cases}$$

The helpfulness of product reviews is modeled by Gain and their scores, higher the rank of review, more the helpful information.

D. Bilinear Model

Similarity learning is another area of supervised machine learning in artificial intelligence. Several attempts have been made to determine the relevance of a query *q* to a document *d*, e.g., similarity of two documents to detect plagiarism. Similarity property must satisfy sim(d,d)=1, sim(d,d_1)=sim(d_1,d), where *d* and d_1 are documents. This model originates from simple cosine similarity or Euclidean dot product

$$cos(q,d) = \frac{q \bullet d}{|q||d|}$$

Closer cos(q,d) is to 1 more similar they are, closer cos(q,d) to zero, more dissimilar they are. Cosine similarity has one problem, that common irrelevant words can dominate the ranking. It is resolved by using the weighted cosine measure [11]

$$cos_w(q,d) = \frac{(q \odot d) \bullet w}{|q||d|}$$

where *w* is the weight vector and ⊙ is Hadamard product operator. This definition violates the similarity property sim(d,d) = 1. Since the weight vector has positive components, we take the point wise positive square root of components, and get w=$\sqrt{w} \odot \sqrt{w}$. Now the accurate weighted cosine is accurately defined as

$$cos_w(q,d) = cos\,(q_w, d_w) = \frac{q_w \bullet d_w}{|q_w||d_w|}$$

where $q_w = q \odot \sqrt{w}$ and $d_w = d \odot \sqrt{w}$

It could not detect words that appear many times in a selected document, but which are rare among other documents. A new metric BM25 [5] was introduced to resolve this.

$$bm25(q,d) = \sum_{i=1}^{n} \frac{IDF(q_i) \bullet f(q_i, d) \bullet (k_1 + 1)}{f(q_i, d) + k_1 \bullet (1 - b + b \bullet \frac{|d|}{avgdl})}$$

where parameters q_i is a word in the query *q*, *d* is a document, *f* is term frequency of q_i, IDF inverse document frequency, '*avgdl'* is the average document length, *b* and k_1 are weight parameters; see [12, 13, 14] for further detail.

The problem still persisted. These could not resolve query and document to be of different lengths and another problem was with the similar words. Bilinear measure uses the concepts in these measures and is able to

overcome their deficiencies. The bilinear model can compute the similarity of objects of different dimensions.

TF-IDF like measures resolve some of the similarity issues: (1) irrelevant words issue of cosine measure via weighted cosine and (2) words frequent/important in one and rare in other documents issue via bm25 measure. But still there *are other issues (1) different length documents, a query, and a document are mostly of different lengths, (2) concerning synonyms, i.e., different words being used to refer to the same concept, e.g.* bright, sharp. *(3) the same word may have different meanings, e.g.* it is my "fan" versus he is my "fan". Such problems are not resolved by the above-mentioned similarity measures. Since queries/reviews are done by different people, it is important to differentiate whether the questions and reviews are tangentially related and may be drawn from very different vocabularies [15], [8]. Thus, one needs to learn whether a word (say burning) is used for "fire" or "heat" in a given scenario. Bilinear models [16, 17] can help to address this issue by learning complex mappings between words in one corpus and words in another (or more generally between arbitrary feature spaces). A common approach for learning similarity is to learn correlation matrix M so that the similarity metric for q and d of possibly different lengths becomes q^TMd. The correlation matrix is also suitable for streaming data. The matrix M can be quickly updated by a simple matrix addition. Thus, the compatibility between a query and a document is given by

$$q^T M d = \sum_{i=1,n;j=1,n} M_{ij}\, q_i\, d_j$$

where M_{ij} encodes the relationship between a term q_i in the query and a term d_j in document d. It incorporates the relation between different words while Cosine handles the relation between the identical lengths query and document. This is a flexible metric because the dimensions of q and d can be different. In practice, M is very high-dimensional (square of the size of the vocabulary) sparse matrix. Base on the size of vocabulary as compared to use in a query or document, we assume that it is a low rank matrix. Any matrix M can be decomposed as $M = USV^T$ where S is a diagonal matrix with a few non-zero diagonal entries (eigenvalues of M), U and V are orthogonal matrices. They can be reduced depending on the rank of M. This is known as singular value decomposition (SVD) of M. For a related article see [18]. Now the similarity measure becomes

$$q^TMd = q^TUSV^Td = q^TUS(d^TV)^T$$

$$= (U^Tq)^TS\,(V^Td)$$

Intuitively, U^Tq and V^Td are projections of q and d on smaller spaces than the query and the document dimensions.

These projections provide useful information about the dominant words. An interesting consequence of projections is that synonyms in q and d are projected to nearby words resulting higher inner product. Also, it optimizes the calculation over q^TMd by replacing M by new M using reduced dimensions of U, S, and V.

Note. In [5] the matrix M is decomposed into A and B such that $M=AB^T$. It does not ensure that A and B are orthogonal matrices. This decomposition is *not unique* because there are several QR type decompositions, for example, we can have A=U, B=SV, or A=US, B=V or $A=U\sqrt{S}$ and $B=\sqrt{S}V$ etc.

It is not beneficial to apply data mining algorithm before preprocessing the data. Bilinear similarity implicitly cleans up the data in resolving several issues mentioned earlier. Now that data has been reduced, we can apply entropy to the dataset with reduced dimensions of q and d. his yields a hybrid algorithm. In the previous section, we used Entropy on raw data. Overall we have hybrid data QA and text, hybrid algorithm using entropy, and bilinear similarity.

In the experiments section, we use entropy and bilinear similarity of q and d. Next step will be to compare the outcomes by using compressed U^Tq and V^Td instead of q and d. The hybrid algorithm will give even better decision power than the application of standalone algorithms.

III. Contributions of this Paper

A. Hybrid Study

This study began with a wider goal to develop novel algorithms, not just selectively use available algorithms, for better online classification and ranking of products and services. We ended up defining a hybrid approach to take advantage of the best features related to review data and available classifiers.

Amazon review data for various products is both solicited and unsolicited. Unsolicited data ends up at the blogs and solicited data is in the form of queries, query by form (QBF). Unsolicited data is naturally unstructured. The QBF data is somewhat structured and is used in our research study to evaluate the effectiveness of the proposed algorithm. The raw data that we have used is available freely for use at [20] and is discussed in detail in sub-section C.

For our study reported in this paper, we are using two types of records. First category of record type is in the form of plain text and contains nine attributes, they are: nominal (titleOfReview, reviewerID, reviewerName, productID, reviewText in the form of text comments), ordinal (productRating 0 to 5, ratingFrequency [helpful ratings, total ratings], interval (unixReviewTime, reviewDate). Here five of the attributes are irrelevant to the product review. The second category of data type is in

the form of questions and answers. Each record contains seven attributes: productID, questionType (y/n or open-ended), answerType, questionText, answerText, unixTime, date. Again three of the attributes are not relevant to the analysis.

We start with two categories of opinion based data, text reviews and QA, and calculate similarity score for data of each type uniformly. We then take the results and use the weighted contribution of each similar to the approach taken in [19]. We propose the weighting of each may be calculated from the datasets as follows. If the r_T is the number of reviews in text form and r_Q is the number of reviews in question/answer form, their contribution is weighted with weight vector $\frac{(r_T, r_Q)}{r_T + r_Q}$.

The value of such a model is relevance, classification, and ranking parameters are learned simultaneously. This allows individual learners to focus on classifying only those instances that are 'relevant,' without considering the irrelevant instances.

B. Ranking Algorithm

We take advantage of many of the techniques surveyed in section II into a single adaptive hybrid algorithm to rank products. We use the following heterogeneous information related to a product as inputs to our approach:

1. Text of the review
2. Question/Answer data
3. Rating of the product
4. Rating of the reviews

The QA data is further divided into the following categories:

1. Normalized ratio of asked questions to answered questions, where 0 would indicate no answers and 1 would indicate all questions are answered.

2. Normalized relevance rating of the answer, where 0 would indicate the answer is not relevant and 1 would indicate the answer is completely relevant.

Overall the product is ranked on weighted QA rank, weighted text based review rank, and normalized rating using the adaptive hybrid algorithm defined in Figure 1. The weighting function of reviews vs QA ranks is calculated from the number of related dataset records. For example, if the r_T is the number of records in text form reviews and r_Q is the number of records in question/answer form, then the weight vector (γ,δ) is calculated as:

$$(\gamma, \delta) = \left(\frac{r_T}{r_T + r_Q}, \frac{r_Q}{r_T + r_Q}\right)$$

Algorithm: A product 'i's final rank is given by:

$$productRank_i = \alpha\,(productRank_i) + \beta\left(\gamma\left(\frac{productRating_{,i}{}^{review_{rank,i}}}{2}\right) + \delta * QA_{rank,i}\right)$$

where,

- α and β are regression coefficients, α+ β=1
- γ and δ are review data size weight factors, γ + δ=1
- $productRating_{,i}$: The normalized over [0,2] average of all product 'i' ratings received from reviewers. The initial value is 1.
- $review_{rank,i} = \begin{Bmatrix} 0.5\ entropyRating_i + 0.5\ \frac{helpfulRating_{count,i}}{totalRating_{count,i}}, & if\ totalRating_{count,i} > 0 \\ entropyRating_i, & otherwise \end{Bmatrix}$
- $QA_{rank,i} = 0.5\ answered_{ratio,i} + 0.5\ bilinearSimilarity_{score,i}$
- $answered_{ratio,i}$: the ratio of answered questions to asked questions
- $bilineaerSimilarity_{score,i}$: normalized relevance rating of the answers using bilinear similarity

Figure 1. Product ranking algorithm

The algorithm has the following desirable properties:

1. The coefficients α and β can be fine-tuned to influence the adaptability of the rank to new information presented (in terms of reviews and Q&A). If the rank is α heavy, i.e., $\alpha > 0.5$, the rank is more stable whereas if it is β heavy, i.e., $\beta > 0.5$, then the algorithm is more reactive to new information
2. $review_{rank}$ is calculated by blending both user helpfulness rating and also the entropy-based machine learning rating, see sections II.C.1 and II.C.2. This removes any implicit or explicit bias in a particular product's reviews and uses the information gathered as the collective wisdom to make the rating more neutral.
3. Helpful reviews fused with higher *productRating* (greater than 1) score improves the overall product rank. Whereas, helpful reviews combined with the lower *productRating* (less than 1) score, pulls it down further. See Fig. 2 for illustration.

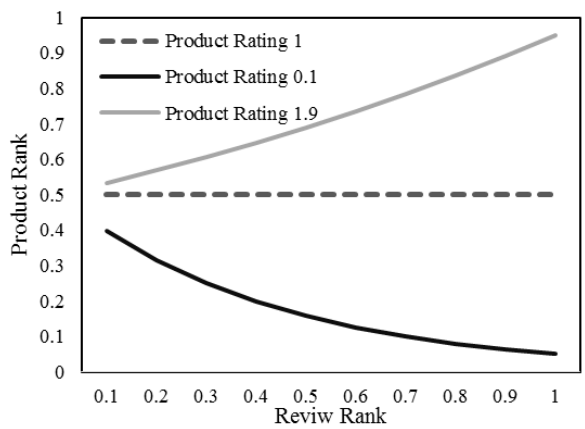


Figure 2. Impact of Review Rank on overall Product Rank for given Product Rating

C. Dataset

We used Amazon product review data for our analysis, particularly, the star ranking, reviews, and related Q&A. The raw data is available freely for use at [20] and contains product reviews and metadata from Amazon, including 142.8 million reviews spanning May 1996 - July 2014. This dataset includes reviews (ratings, text, helpfulness votes), product metadata (descriptions, category information, price, brand, and image features), and links (also viewed/also bought graphs).and has been used in related research, such as [21].

As mentioned above, we are particularly interested in star ranking, review text, and Q&A text. For our analysis, we utilized the k-core subsets of Musical Instruments (500,000 records), Electronics (800,000 records), and Health and Personal Care (900,000 records) categories. Sample records are presented below:

{
"reviewerID": "A3LA5EHF2WWNFK",
"asin": "B000000545",
"reviewerName": "Philip Y.",
"helpful": [1, 1],
"reviewText": "One of the best albums I've ever heard. B-Legit's great delivery and lyrics along with many fly guest featured over this tight production - it just makes a perfect album with the bay area flava. C-Bo, Kurupt, E40, Celly Cell and many more join the Savage in one of the west coast best and most underrated albums. Get that fly sh*t!!!",
"overall": 5.0,
"summary": "OFF THE HEEZY!!!!",
"unixReviewTime": 906681600,
"reviewTime": "09 25, 1998"
}

where

- *reviewerID* - ID of the reviewer
- *asin* - ID of the product
- *reviewerName* - name of the reviewer
- *helpful* - helpfulness rating of the review as [helpful ratings, total ratings]
- *reviewText* - text of the review
- *overall* - rating of the product out of [0, 5]
- *summary* – summary/title of the review
- *unixReviewTime* - time of the review (unix time)
- *reviewTime* - time of the review (raw)

Review data record in the form of QA:

{
'questionType': 'yes/no',
'asin': '9792372326',
'answerTime': 'Jan 21, 2015',
'unixTime': 1421827200,
'question': 'Will the K10 be damaged if its XLR output is connected to a mixing board and somebody accidentally pushes the +48V phantom power button on the board?',
'answerType': 'N',
'answer': 'no'
}

where

- *questionType* - type of question. Could be 'yes/no' or 'open-ended'
- *asin* - ID of the product
- *answerType* - type of answer. Could be 'Y', 'N', or '?' (if the polarity of the answer could not be predicted). Only present for yes/no questions.
- *answerTime* - raw answer timestamp
- *unixTime* - answer timestamp converted to unix time
- *question* - question text
- *answer* - answer text

IV. Evaluation and Discussion

We start by calculating the entropy measure of product reviews. Using the entropy based classification model as explained in section II B, we calculate the $Score_i$ for the reviews per product '*i*'. We have used k-fold cross validation, k=10, to optimize the model parameters to make the model fit the training data as best as possible.

We have used Amazon's Elastic Map Reduce (EMR) service, using Java, to sift through both the review and the QA files, see description of dataset in the next section. Below is the brief description of the two phases used, also depicted in Figure 3.

Phase 1: The mappers map the review and QA data to a product ASIN, whereas the Reducers output the data per product in the form of a string array.

Phase 2: The mappers calculate the stats, entropy and bilinearity, per product, whereas the Reducers take all the relevant stats per product and using algorithm given in Figure 1, output the final rank.

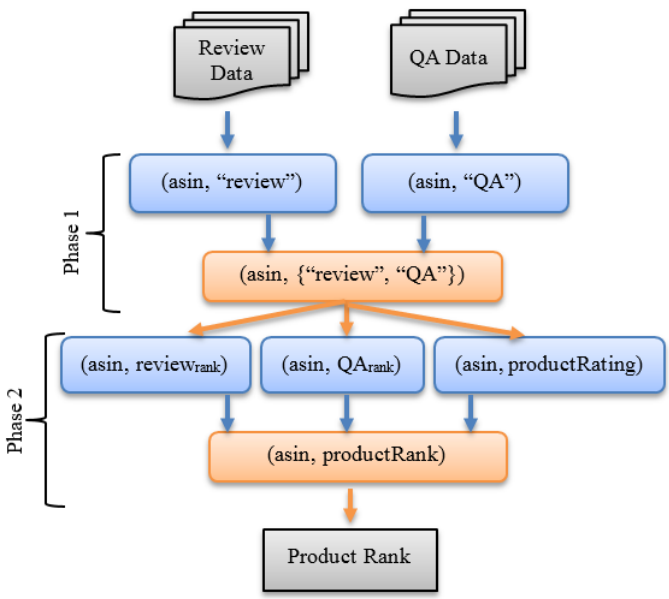


Figure 3. MapReduce Phases for calculating the Product Rank. Blue boxes represent Mappers and orange boxes represent Reducers.

The $entropyRating_i$ measure, as used in the algorithm presented above, is calculated as:

$$entropyRating_i = \frac{Score_i}{Score_{max}}$$

$$Score_{max} = \max \forall Score_i$$

Hence entropyRating will be 1 for the product with maximum score, and will be distributed in the unit interval [0, 1] for all other products.

Table I and II present the precision, recall and F-measures for the entropy-based classifications of reviews as 'helpful' and 'unhelpful'

TABLE I. Precision, Recall, and F-Measure for Reviews Classified as 'Helpful'

	Precision	Recall	F-Measure
Musical Instruments	0.78	0.76	0.76987
Health and Personal Care	0.76	0.77	0.764967
Electronics	0.72	0.79	0.753377

TABLE II. Precision, Recall and F-Measure for Reviews Classified as 'Not Helpful'

	Precision	Recall	F-Measure
Musical Instruments	0.79	0.75	0.769480519
Health and Personal Care	0.75	0.78	0.764705882
Electronics	0.74	0.77	0.754701987

In Table III we present some product ranking examples. The product rank is calculated by both the old method, which is simply the average of all the product ratings received by the users and the new method, which is based on our algorithm. We would like to elaborate on the interesting contrasts in the two methods. If you look at examples 1-3, the old ranking is the same for all three cases, as the product rating is the same. However, the new ranking is quite different for the three cases. It is all the more interesting is to see how the new product ranking accurately reflects the review ranks and QA ranks, with higher values reflected in higher ranks and vice versa.

TABLE III. Product Rank as Calculated by Old and New Algorithm

	Product Rating	Review Rank	QA Rank	Product rank (new)	Product rank (old)
1	1.9	0.5	0.5	0.5946012	0.95
2	1.9	0.8	0.8	0.8177754	0.95
3	1.9	0.3	0.3	0.4530861	0.95
4	1.1	0.8	0.8	0.6698076	0.55
5	1.1	0.8	0.4	0.4698076	0.55
6	0.5	0.8	0.8	0.5435873	0.25
7	0.5	0.3	0.8	0.6030631	0.25

Examples 4 and 5 show the effect of different QA rank on the overall product ranking where the product rating and review rank is the same. Where the old product rank is

solely dependent on product rating, the new rank reflects the QA ranking proportionally. Similarly, examples 6 and 7 reflect the effects of different review rank when product rating and QA rank are the same. It is, however, important to note that here the higher review rank resulted in lower overall product rank. This is one of the key characteristics of our algorithm, i.e., helpful reviews on a product increase the authenticity of the lower rating. Hence, two products with same low product rating but one having more 'helpful' reviews is an indication of a strong negative response towards the product. Hence in the overall ranking, this product falls further below the rank list.

V. Conclusion and Future Work

Several commercial websites for products and services provide platforms for the consumers to share their opinions. In this paper, we developed a hybrid adaptive algorithm that uses heterogeneous product review data (star ranking, text based reviews, question/answer data) to simultaneously analyze and evaluate the reviews in order to rank similar products. We show that adapting the ranking algorithm by simultaneously using Entropy and Bilinear Similarity measures yields more accurate evaluation than using them in isolation and on restricted data types. For experimentation and evaluation we used three categories of Amazon.com review data: k-core subsets of Musical Instruments, Health and Personal Care and Electronics to simultaneously analyze and evaluate the reviews for product ranking. The results of experiments are displayed using Gold standard metrics with the help of tables. We showed that our hybrid approach is effective and yields results which make not only intuitive sense but are mathematically sound.

In this work we have used weighted scores of Entropy and Bilinearity breadth first techniques. In future, we would like to explore the potential for depth first hybrid algorithm. Also, we would like to further expand this work to enable product related predictions for both consumers and sellers.

Acknowledgment

Use of Amazon product review data, available freely for research [20]

References

[1] G. Lackermair, D. Kailer and K. Kanmaz, "Importance of Online Product Reviews from a Consumer's Perspective". Advances in Economics and Business Vol. 1(1) 2013. Pages 1-5

[2] P. D. Turney, "Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews". Proceeding ACL 2002 Proceedings of the 40th Annual Meeting on Association for Computational Linguistics. Pages 417-424

[3] Aria Haghighi, How to approach machine learning as a non-technical person. Crunch Network, Apr 2, 2016. http://techcrunch.com/2016/04/02/how-to-approach-machine-learning-as-a-non-technical-person/

[4] D.H. Park, J. Lee, and I. Han, "The Effect of On-Line Consumer Reviews on Consumer Purchasing Intention: The Moderating Role of Involvement". International Journal of Electronic Commerce Vol. 11, No. 4 (Summer, 2007). Pages 125-148

[5] L. Zhang, D. Agarwal and B. C. Chen, "Generalizing matrix factorization through flexible regression priors". In Proceedings of the fifth ACM conference on Recommender systems 2011, Pages 13-20

[6] H. Yu and V. Hatzivassiloglou. "Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences". Proceeding of EMNLP '03 Proceedings of the 2003 conference on Empirical methods in natural language processing. Pages 129-136

[7] D. H. Park, H. D. Kim, C. Zhai and L. Guo, "Retrieval of Relevant Opinion Sentences for New Products". SIGIR 2015 Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval. Pages 393-402

[8] R. Zhang and T. Thomas, "An Entropy-Based Model for Discovering the Usefulness of Online Product Reviews". In Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web. Pages 759-762

[9] M. Hu and B. Liu, "Mining and summarizing customer reviews". KDD 2004 Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. Pages 168-177

[10] C. E. Shannon and W. Weaver, "The Mathematical Theory of Communication". University of Illionoins Press

[11] J. McAuley and A. Yang, "Addressing Complex and Subjective Product-Related Queries with Customer Reviews". WWW 2016 Proceedings of the 25th International Conference on World Wide Web. Pages 625-635

[12] K. S. Jones, S. Walker and S. Robertson, "A probabilistic model of information retrieval: development and comparative experiments". Journal of Information Processing and Management: an International Journal, Volume 36 Issue 6, 2000. Pages 779-808

[13] J. Jeon, W. B. Croft and J. H. Lee. "Finding similar questions in large question and answer archives". CIKM 2005 Proceedings of the 14th ACM international conference on Information and knowledge management. Pages 84-90

[14] C. D. Manning, P. Raghavan, and H. Schültze, An Introduction to Information Retrieval. Cambridge University Press, 2009.

[15] A. Berger, R. Caruana, D. Cohn, D. Freitag, and V. Mittal, "Bridging the lexical chasm: statistical approaches to answer finding". SIGIR 2000 Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval. Pages 192-199

[16] W. Chu and S.-T. Park, "Personalized recommendation on dynamic content using predictive bilinear models". WWW 2009 Proceedings of the 18th international conference on World wide web. Pages 691-700

[17] W. T. Freeman and J. Tenenbaum, "Learning bilinear models for two-factor problems in vision". IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1997. Pages 554-560

[18] C. L. Sabharwal and B. Anjum, "Principal Component Analysis as an Integral Part of Data Mining in Health Informatics". Proceedings of 31st International Conference on Computers and Their Applications, CATA 2016. Pages 251-256.

[19] R. A. Jacobs, M. I. Jordan, S. J. Nowlan and G. E. Hinton, "Adaptive mixtures of local experts". Journal of Neural Computation. Volume 3 Issue 1, Spring 1991. Pages 79-87

[20] Amazon Review Data: jmcauley.ucsd.edu/data/amazon

[21] J. McAuley, R. Pandey and J. Leskovec, "Inferring Networks of Substitutable and Complementary Products". KDD 2015 Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Pages 785-794

Novel Group Detection and Analysis Method Based on Automatic and Fast Density Clustering

Ke Jin, WeiWei Xing
School of Software Engineering
Beijing Jiaotong University
Beijing 100044, China
wwxing@bjtu.edu.cn

Peng Bao
School of Software Engineering
Beijing Jiaotong University
Beijing 100044, China

***Abstract*—Group-level crowd behavior analysis is a new and promising method with important applications for the video surveillance and understanding of crowds. However, a specific definition for a group in a crowd field has rarely been investigated. This paper proposes a complete and reasonable definition for a group in a crowd field and presents a fast and automatic group detection method. First, automatic and fast density clustering (AFDC) is used to find the group core, which is then refined based on the property of coherent neighbor invariance. This detection method is more adaptive to groups with arbitrary shapes and varying densities because the group core is refined with coherent neighbors. Experiments on hundreds of video clips of public scenes showed that the method achieved an excellent detection performance and attractive statistical results. In particular, the number of people in a group exhibits a power-law distribution truncated by an exponential tail; this is significant to understanding crowd scenes and crowd simulation.**

Keywords—group detection; group level; crowd behavior analysis; density clustering; power-law distribution

I. INTRODUCTION

A crowd is a large number of persons gathered closely together. The behavior analysis of crowds has important applications in public areas of high security interest. Because groups are the primary entities that make up a crowd and impact crowd dynamics [1, 2], group-level behavior analysis is a very important method for characterizing and comparing the dynamics of different crowded scenes. Understanding group dynamics is scientifically important and practically useful for crowd behavior analysis [3, 4]. However, a specific definition for a group in a crowd has rarely been given, even though groups have been extensively studied in the socio−psychological and biological fields [5–7].

There is a rich body of literature [1, 8–10] suggesting that a crowd generally maintains the important property of collectiveness, i.e., the majority of pedestrians tends to move in groups with friends and family members. The state of a group can be greatly different: it may be stationary at some point but become mobile following crowd dynamics with diverse densities and varying shapes other times. The collective phenomenon is more prominent in dense crowds [2, 11]. As a set of members with a common goal and collective behavior, a group often exhibits two other properties in the real world: overlapping and hierarchy. These can be abstracted from the common phenomenon where some individuals of a group move towards the same goal while being surrounded by other coherent individuals. This is important when analyzing crowd behavior at the group level. Fig. 1 shows some examples. Overlapping means that some individuals belonging to one group may be part of another group at another time, especially when the group

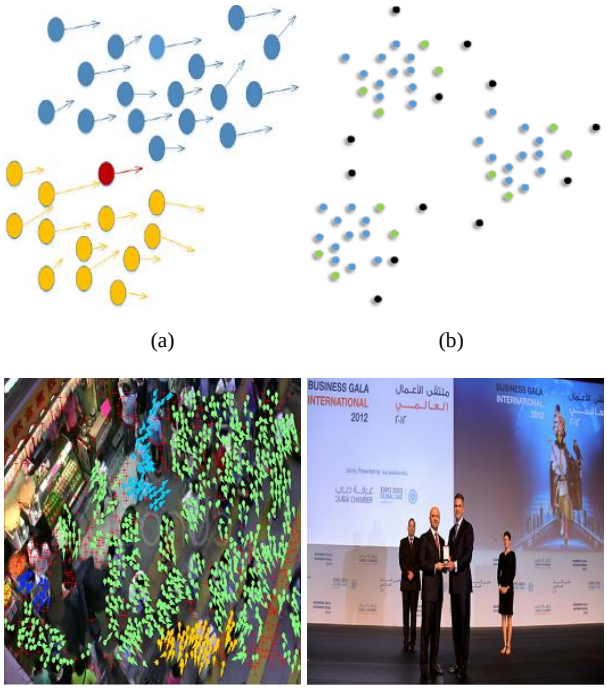


Figure 1. Examples of overlapping and hierarchical properties of a group. (a) The red individual was originally part of the blue group and now belongs to the orange group. (b) The group can be viewed as hierarchical: the stable individuals comprise the group core (blue), which collects other coherent individuals (green) into a group. The remaining individuals not belongs to any group are outliers (black). (c) The collectiveness of a group in the real world; the two groups in opposite directions merge into a bigger group. (d) The important people (person giving the award and person receiving it) may be part of the group core (all four people).

DOI reference number: 10.18293/DMS2016-019

splits or merges in extremely crowded situations. Depending on the level of importance and correlation with the group center, individuals in a group can be divided into two levels: the group core and coherent individuals. Thus, groups can be viewed as hierarchical: stable individuals comprise the group core, which collects other coherent individuals into a group.

To better understand these properties of a group in a crowd, the problem of detecting a group must be solved first. Many methods have been proposed to detect groups [1, 12–14]. However, these methods cannot be applied to complex patterns. Later works [6, 15] have developed models by abstracting the spatiotemporal coherent relationships to provide more satisfying results. These emphasized finding the group without accounting for the hierarchical attributes.

Group detection can be a complex problem in crowded scenes [16], especially when considering groups with arbitrary shapes and varying densities. However, hierarchical clustering (HC) [8] and the density peaks method [17] cluster tracklets under the assumption that the cluster center is surrounded by neighbors at a lower density or by using velocity and spatial constraints. It is difficult for these methods to handle crowd problems with varying densities. Other methods [18, 19] that use a motion flow field are limited to specifying the clustering number manually. The most recent methods [17, 7] require a factor given manually, which needs field experience. Shuai et al. [2, 11] tried to analyze group attributes but could only handle the problem when the groups are stationary.

For application to video surveillance, another key challenge is efficiency. The diverse scenes from large amounts of video require a much more efficient way to quickly detect group dynamics. Approaches based on optimal flow [4, 13, 20, 21] consider large amounts of tracking points, and methods based on local coherent relationships [3, 6, 8, 15] need to maintain the topological structure and update it with iterative computations. The time cost for all of these is so high that they cannot be used in real-time applications.

The goal of our research was to profile a group in a crowd, give a specific visual definition, detect the group, study the statistical properties, and consider the impact on crowd dynamics. We considered a group to be a set of members with a common goal and collective behavior and having two common properties: overlapping and hierarchy. Groups are the primary entities that make up a crowd. Their dynamics are influenced by the crowd dynamics, and their impact on crowd cannot be ignored either. To the best of our knowledge, we are the first to try to comprehensively and systematically give a relatively specific definition based on previous research. The main contributions of our approach include the following:

- A novel automatic and fast density clustering algorithm is proposed. There are three salient properties must not be ignored: (a) recognizing outliers accurately, (b) automatically determining the group number and group center without any manual intervention, and (c) accurate, efficient, and easy real-time application.
- A new fast and robust group detection method is developed that is based on the Automatic and Fast Density Clustering (AFDC) algorithm and combined with spatiotemporal neighbor invariance to characterize the underlying priors between neighbors. This detection method is more robust against complex scenes and it is also presented to uncover both local and global attributes. It can be effectively applied to groups with arbitrary shapes and varying densities.
- A short but comprehensive definition is presented for depicting groups in a crowd along with interesting statistical findings. Since it has attracted a great number of attention, there is no proper definition for a group. Based on a great deal of research, we limited a “group” to a specific meaning in a crowd field. This served as an inspiration for group understanding and crowd behavior simulation. Similar to collective phenomena in colonies [5], human groups also show interesting statistical results.

II. Related Works

For a long time, crowd analysis has been conducted at the macroscopic or microscopic level. At the macroscopic level, the main research focus has been the global motions of massive numbers of people; crowd behavior analysis did not consider the movements of individuals [4, 22–24]. At the microscopic level, the crowd is treated as a collection of individuals [25, 26], and analysis is based on the collective information of individuals. In contrast, some studies tried to analyze crowds at the group level [3, 8]. There are two main methods for recognizing groups in a crowd: shape classification and trajectory analysis. There are no fine methods for group detection by shape classification, while methods based on trajectory/tracklets greatly rely on object detection and tracking.

Coherent motion is an important phenomenon that has attracted a great deal of attention in a wide range of fields, from biology [5, 27] to physics [28], and cannot be ignored in video analysis. There are many similarities between group dynamics and coherent motion; both have the same goal and demonstrate collectiveness in a crowd.

State-of-the-art methods [3, 6, 7] have achieved group detection through tracklet analysis. Shao et al. [3] proposed the novel Collective Transition (CT) prior to capture the underlying dynamics of a group and devised a set of visual descriptors to quantify the universal properties of groups in a crowd. Wu et al. [7] developed the Collective Density Clustering (CDC) approach, and Zhou et al. [6] presented Coherent Filtering (CF) to recognize coherent motion. However, the above methods are unfeasible for real-time application or require manual intervention. More importantly, none of them analyze crowd behavior at the group level nor learn groups with both overlapping and hierarchical properties.

Other works [23, 29] focusing on scene-specific analysis may not be applied to groups with arbitrary shapes and varying densities. Our work is significantly different from the above research as follows: (i) We analyzed crowd behaviors at the group level and considered the overlapping and hierarchical properties when profiling a group, (ii) and our approach can handle groups both in crowded scenes and in low-density or disconnected areas.

III. GROUP DETECTION METHOD

We define a group as a collection of individuals who gather together and move towards the same goal. Examples in the real world have shown us that groups are usually overlapping and hierarchical. A group is generally a special part of crowd but goes beyond a simple collection of several individuals. The members of a group are usually spatially proximate or temporally coherent, and a group is a dynamic and independent unit that demonstrates various fundamental properties. Its activities can greatly influence the crowd in diverse modes.

In this paper, we propose a group detection method based on this definition. First, we detect the group core automatically by quickly finding the peak density of individuals. The accuracy and efficiency of our method guarantee its effectiveness for real-time application and make it robust against arbitrary shapes and varying densities in changing scenes. We then refine the group members with the coherent neighbor invariance prior. This reflects the overlapping property of the group and our approach achieves detection with both local and global properties. Groups vary in different scenes. Examples are shown in Fig. 2.

Motion features are detected with the gKLT feature point tracker [15]. Each group $\{G_i\}$ encompasses a set of tracklets $\{z\}$. For better performance, the expectation–maximization (EM) algorithm is used to fit the tracklets.

A. *Automatic and Fast Density Clustering*

The AFDC algorithm is the main part of the group detection method. First, we set the parameters for the algorithm, and we explain the differences between our algorithm and the normal one [17]. We then find the group center for each group and solve for the group members. Here, the group members comprise the group core.

1) *Initial Parameters*

In contrast to the normal density calculation method [17], our method can estimate both the position and orientation to solve the problem.

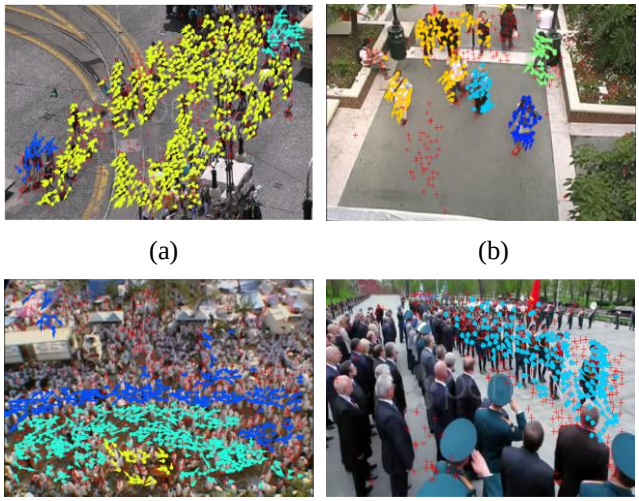
In two-dimensional space, consider a crowd $\{C\}$ possessing N points, where $\{C_i\}$ is the set of points and $\|C\| = N$. We consider every point C_i $(i = 1 \dots N)$ as a vector with the spatial position (x_i, y_i) and orientation α_i. The orientation for each point at the certain frame can be calculated by neighbouring several frames. The similarities (including distance and orientation) between two points C_i and C_j are denoted as

$$d_{xy}(i,j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{1}$$

$$\theta(i,j) = |\alpha_i - \alpha_j| \tag{2}$$

Now, we can calculate the density ρ_i as a Gaussian kernel function:

$$\rho_i = \sum_j \exp\left(-\frac{d_{xy}(i,j) * (\cos\theta(i,j)+1)}{d_{cut}}\right) \tag{3}$$



(a) (b)

(b) (d)

Figure 2. Groups in several representative crowd scenes. Scatters and arrows with different colors indicate different groups, and red crosses indicate the outliers. (a) Non-uniform group. (b) Crowd with low density. (c) Groups in an extremely crowded scene. (d) Marching soldiers and stationary nearby visitors.

The salient difference between this parameter and the primary one [17] is that the position and orientation can be synthetically estimated at the same time. As we know, $\cos\theta$ is a monotonic decreasing function from 1 to 0, where the independent variable θ changes from 0 to $\frac{\pi}{2}$. Thus, the orientation θ can adjust the error grouping for only the distance, but a small θ will be in the same group. In addition, δ_i represents the closest distance from other points with a higher density ρ_i. For the point C_i, δ_i is denoted as

$$\delta_i = min_{j:\, \rho_j > \rho_i}\, d_{xy}\,(i,j) \tag{4}$$

For the point with the highest density, $\delta_i = \max d_{xy}\,(i,j)$.

Because the accuracy of ρ_i heavily depends on the threshold d_{cut}, we propose an automatic method to handle this inevitable problem by using entropy without depending on subjective experience [7, 17]. We were inspired by the Shannon entropy; in information theory, entropy is often used to quantify the uncertainty of a probability distribution when a specific event is observed. We calculate the entropy H for the point C_i $(i = 1 \dots N)$ with the density δ_i as follows:

$$H = -\sum_i^N \frac{\rho_i}{W} \log \frac{\rho_i}{W} \tag{5}$$

where $W = \sum_i^N \rho_i$ is the normalization factor. We then establish the function H with the independent variable d_{cut}. We performed experiments to prove the effectiveness of this method.

2) *Detecting the Group Core*

Inspired by the new clustering method [17], the first step after the density ρ_i and distance δ_i are defined is determining the group center for each subgroup. A larger value of ρ_i indicates that the point C_i has more neighboring points, and δ_i is much

Algorithm 1 Automatic and Fast Density Clustering

Input: Number of points $\boldsymbol{N}$, set of points $\{C_i\}$ $(i = 1 \dots N)$

Output: Number of group $\boldsymbol{M}$, the clustering result $\{Group_j \mid (j = 1 \dots M)\}$

FOR $i = 1$ **TO** $\boldsymbol{N}$
  FOR j = i+1 **TO** $\boldsymbol{N}$
    $d_{xy}(i,j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$
    $\theta(i,j) = |\alpha_i - \alpha_j|$
  END
END
FOR $i = 1$ **TO** $\boldsymbol{N}$
  $\rho_i = \sum_j \exp\left(-\frac{d_{xy}(i,j) * (\cos\theta(i,j)+1)}{d_{cut}}\right)$
END
FOR $i = 1$ **TO** $\boldsymbol{N}$
  IF ρ_i is larger than any other points
    $\delta_i = \max d_{xy}\ (i,j)$
  ELSE $\delta_i = min_{j:\rho_j > \rho_i}\ d_{xy}\ (i,j)$
END
FOR $i = 1$ **TO** $\boldsymbol{N}$
  $product_i = \rho_i * \delta_i$
END
SORT$(product_i, DESC)$/*sort the $product_i$*/
FOR $i = 2$ **TO** $\boldsymbol{N}$
  $gap_i = product_{i-1} - product_i$
  IF gap_i > **TH**
    $Group_i$ = i-1
  ELSE BREAK
END
FOR $i = 1$ **TO** $\boldsymbol{N}$
  Assign point according its distance to group center.
END
Return: $\{Group_j \mid (j = 1 \dots M)\}$

larger than the nearest neighbor distance only when C_i has the maximum density. We determine the group centers subjectively according to the larger ρ_i and δ_i. In contrast to the typical method in [17], the initial group center is determined by $max_i(\rho_i \times \delta_i)$. We select the value for which the gap between the neighboring two $(\rho_i \times \delta_i)$ is larger than a threshold TH, which is far above the other gap. $Group_i = i$ is used to number the group centers. Other group members are assigned later in descending order of ρ_i. Except for the initial group centers, other points are assigned as follows. For the points C_i and C_j with $\rho_i < \rho_j$, the point C_i is assigned to the nearest point C_j when $d_{xy} < \frac{d_{cut}}{\cos\theta\ (i,j)+1}$. Our method has two features: (a) points should in the same group with the nearest neighbor of higher density, and (b) points in one group show no great difference in orientation. Thus, it can be used to determine the number of groups.

After we define the center and core of groups, we can also determine outliers, including those inside groups. Outliers often occur when the group contains too few points or the orientations of points in one group are inconsistent. The outliers are grouped into a single point. The method can correctly identify outliers.

Algorithm 1 presents the details of the AFDC algorithm.

B. *Group Refinement with Coherent Neighbor Invariance*

We considered the collectiveness of a group combined with the spatiotemporal relationship between individuals to refine the group core. Experiments on synthetic data have shown the existence of Coherent Neighbor Invariance, including both spatiotemporal and velocity relationships [6].

Based on the underlying prior in the dynamics of crowd, we tried to finding volatile individuals with coherent behavior mainly from observing the time series for group cores. For every point C_i $i = 1 \dots N)$ at time t, where the frame t contains N individuals, we denote the K nearest neighbors as N_t^i. From the time t to $t + d$, the invariant neighbor set $M_{t\to d}^i$ contains the invariant neighbors among the K nearest neighbor set. For a certain individual A in the same group as another individual B, A tends to keep in the KNN set of B [6] over time. We define the new group G as the union of $M_{t\to d}^i$ and the group core, where i indicates the group center individual.

IV. Experimental Results

We evaluated our proposed group detection algorithm for diverse and varying crowd scenes in the real world, such as escalators, crossroads, marathons, and railway stations. Some fascinating group characteristics were observed in the statistical experiments, and the application to crowd simulation demonstrated the effectiveness of our method.

A. *Experimental Setup*

Evaluations were mainly conducted on the CUHK Crowd Dataset for group detection and statistical analysis. This dataset contains crowd videos with various densities and perspective scales, including 474 video clips from 215 scenes. Shao et al. [3] manually annotated tracklets from 300 video clips. Tracklets not belonging to any group were annotated as outliers.

B. *Parameter Setup*

We used the entropy method to automatically determine the only parameter of the AFDC algorithm. We confirmed the effectiveness of the entropy method by designing an experiment with the CUHK Crowd Dataset [3] to investigate how the density threshold influences the quality of the summaries. We tuned the density threshold by varying k from 1% to 5% of the

total N. We calculated d_{cut} with k and N from the experiment and chose the best parameter, as we did before. In general, a large number of experiments should be done to obtain the parameter. Fig. 4 indicates that our group detection method performed the best for specific scenes where the density threshold d_{cut} was set to 60 around $k = 2\%$ and started to drop significantly before 50 and after 80. This proved that our method can detect groups automatically. The density threshold d_{cut} is calculated automatically by choosing the independent variabled_{cut} of the function H when the dependent variable H mutates. Fig. 3 shows an example.

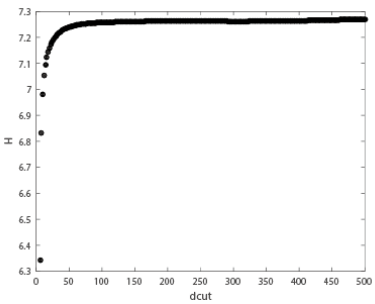


Figure 3. Function H with varying d_{cut}. We select the right d_{cut} at the inflection point.

C. Group Detection Performance

We compared our method with state-of-the-art group detection approaches using Collective Transition priors [3] and Coherent Filtering (CF) [6]. Owing to the expensive time cost of handling video clips of various lengths, we only took the first 30 frames from each clip to analyze. Experiments have shown that considering longer frames does not make a significant difference [3]. The ground truth was based on the CUHK dataset [3]. Fig. 5 compares examples of the detected results and ground truth.

CF is sensitive to tracking errors because it detects coherent motions just by considering neighbor relationships, not the whole dynamics of the whole group. Another reason is that CF first detects coherent motions among neighbors in several frames that it associates with the group. Therefore, its errors accumulate. CF works well in crowded scenes but not in situations when the group merges or splits. CT is a state-of-the art group detection method with novel universal properties for groups in crowds. It works well in both sparse and crowded scenes. However, it cannot identify outliers well, especially those within groups. Because of the preprocessing method needing to maintain a neighbor set and the later iterative process, CT is extremely time-consuming and inapplicable to scenes in real-time. Table I compares the efficiencies of CT and our method in detail. We chose five representative scenes with different frames, sizes, and tracking points. Both CT and our method are based on the gKLT [30] tracking strategy. The efficiency is related to the number of tracking points. The experiments showed that our method works more than ten times faster than CT and can easily be applied to real time scenes.

To some extent, the group detection method can be seen as a clustering process, which is convenient for quantitative evaluation. We adopted three widely used measurements for evaluation: Normalized Mutual Information (NMI) [31], Purity [32], and the Rand Index (RI) [33]. Table II compares the values. Our method and CT performed at the same level, but our method was much more efficient than CT.

(a) $k = 1\%$ (b) $k = 2\%$ (c) $k = 3\%$ (d) $k = 4\%$ (e) $k = 5\%$

Figure 4. Crowd scene: Station. k changes from 1% to 5% of total N

TABLE I. EFFICIENCY COMPARISON WITH CT AND OUR DETECTION METHOD

Methods				**CT**	**Ours**
	Frames	**Tracking points**	**Size**	***Time per frame***	***Time per frame***
Scene 1	184	1500	856 × 480	20.713 s	1.189 s
Scene 2	152	1000	856 × 480	13.564 s	0.649 s
Scene 3	201	1000	1920 × 1080	37.687 s	1.463 s
Scene 4	1001	500	960 × 540	31.498 s	0.501 s
Scene 5	123	1500	1920 × 1080	39.613 s	1.764 s

Figure 5. Comparative results of group detection with three methods in five diverse crowd scenes. Groups are distinguished with colors. Arrows indicate moving directions, and red crosses indicate the outliers. Best viewed in color.

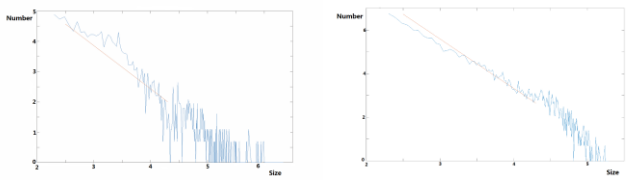
D. Group Statistical Characteristics

Inspired by the collective motion of bacterial colonies [5], we observed some interesting results that will play an important role in crowd simulation and crowd scene understanding. Our statistics for a large number of groups showed that the number of individuals in a group exhibits a power-law distribution truncated by an exponential tail. The statistical properties become more obvious with enough groups. This can be a great guide for crowd simulation and understanding scenes.

With regard to the degree of the power-law distribution, this a common phenomenon, especially for human beings. This is a distinctive property of complex networks and is of interest in activities involving human beings. Fig. 6 shows an example.

TABLE II. COMPARISON OF GROUP DETECTION METHODS

Methods	NMI	Purity	RI
CF [6]	0.43	0.72	0.76
CT [3]	0.49	0.78	0.82
HC [8]	0.27	0.62	0.73
Ours	0.52	0.83	0.86



(a) $N = 2843$ (b) $N = 8547$

Figure 6. The abscissa is the group size, and the ordinate is the number of groups. (a) The total group number N is 2843. (b) The total group number N is 8547. The red straight line fits the power-law distribution, and an exponential tail can be observed.

V. CONCLUSION

In this paper, we systematically analyzed group behavior. used the AFDC algorithm to detect groups in a crowd. We obtained important and interesting attributes intrinsic to a group through a statistical approach. These findings are vital to group

state analysis, crowd simulation, and crowd motion understanding. The advantage of our approach is that it can work with human intervention and provides excellent performance in terms of both accuracy and efficiency. The much higher efficiency of our method due to algorithm optimization and more efficient programming language means that it can be applied in real time to extremely complex scenes.

This work may open an interesting field. Further research is needed, such as video surveillance to define and detect group events and characterizing people to divide a group into three level. In the future, the following improvements will be made to the method: developing a more robust and accurate multi-object detection algorithm for tracklets, developing a better preprocessing algorithm for the tracklets, and reducing the repetitive computation for a faster AFDC algorithm

ACKNOWLEDGMENT

This work is supported in part by National Natural Science Foundation of China (No. 61100143, 61272353, 61370128), Program for New Century Excellent Talents in University (NCET-13-0659), Beijing Higher Education Young Elite Teacher Project (YETP0583), and Fundamental Research Funds for the Central Universities (2014JBZ004, 2015RC031).

REFERENCES

[1] M. Moussaïd, N. Perozo, S. Garnier, D. Helbing, and G. Theraulaz, "The walking behaviour of pedestrian social groups and its impact on crowd dynamics," *PloS One*, vol. 5, no. 4, e10047, 2010.

[2] S. Yi and X. Wang, "Profiling stationary crowd groups," in *IEEE Int. Conf. Multimedia and Expo*, Chengdu, China, Jul. 14-18, 2014, pp. 1–6.

[3] J. Shao, C. C. Loy, and X. Wang, "Scene-independent group profiling in crowd," in *IEEE Conf. Computer Vision and Pattern Recognition*, Columbus, OH, Jun. 23–28, 2014, pp. 2227–2234.

[4] S. Ali and M. Shah, "A Lagrangian particle dynamics approach for crowd flow segmentation and stability analysis," in *IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, Minneapolis, MN, Jun. 18–23, 2007, pp. 1–6

[5] H. P. Zhang, A. Beer, E.-L. Florin, and H. L. Swinney, "Collective motion and density fluctuations in bacterial colonies," *Proc. Natl. Acad. Sci.*, Vol. 107, No. 31, pp. 13626–13630, 2010.

[6] B. Zhou, X. Tang, and X. Wang, "Coherent filtering: Detecting coherent motions from crowd clutters," in *Proc. 12th European Conf. Computer Vision*, Florence, Italy, Oct. 7–13, 2012, part II, pp. 857–871.

[7] Y. Wu, Y. Ye, and C. Zhao, "Coherent motion detection with collective density clustering, in *Proc. 23rd Annu. ACM Conf. Multimedia*, Brisbane, Australia, Oct. 26–30, 2015, pp. 361–370.

[8] W. Ge, R. T. Collins, and B. Ruback, "Vision-based analysis of small groups in pedestrian crowds," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 5, pp. 1003–1016, 2012.

[9] A. F. Aveni, "The not-so-lonely crowd: Friendship groups in collective behavior," *Sociometry*, vol. 40, no. 1, pp. 96–99, 1977.

[10] F. Solera, S. Calderara, and R. Cucchiara, "Structured learning for detection of social groups in crowd," in *10th IEEE Int. Conf. Advanced Video and Signal-based Surveillance*, Krakow, Poland, August 27–30, 2013, pp. 7–12.

[11] S. Yi, X. Wang, C. Lu, and J. Jia, "L0 regularized stationary time estimation for crowd group analysis," in *2014 IEEE Conf. Computer Vision and Pattern Recognition*, Columbus, OH, Jun. 23–28, 2014, pp. 2219–2226.

[12] T. Lan, Y. Wang, W. Yang, S. N. Robinovitch, and G. Mori, "Discriminative latent models for recognizing contextual group activities," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 34, no. 8, pp. 1549–1562, 2012.

[13] L. Kratz and K. Nishino, "Going with the flow: Pedestrian efficiency in crowded scenes," in *Proc. 12th European Conf. Computer Vision*, Florence, Italy, Oct. 7–13, 2012, part IV, pp. 558–572.

[14] C. Zhang, X. Yang, J. Zhu, and W. Lin, "Parsing collective behaviors by hierarchical model with varying structure," in *Proc. 20th ACM Multimedia Conf.*, Nara, Japan, Oct. 29–Nov. 2, 2012, pp. 1085–1088.

[15] B. Zhou, X. Tang, H. Zhang, and X. Wang, "Measuring crowd collectiveness," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 36, no. 8, 1586-1599, 2014.

[16] T. Li, H. Chang, M. Wang, B. Ni, R. Hong, and S. Yan, "Crowded scene analysis: A survey," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 3, pp. 367–386, 2015.

[17] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science*, vol. 344, no. 6191, pp. 1492–1496, 2014.

[18] T. Brox and J. Malik, "Object segmentation by long term analysis of point trajectories," in *Proc. 11th European Conf. Computer Vision*, Heraklion, Greece, Sep. 5-11, 2010, part V, pp. 282–295.

[19] M. Hu, S. Ali, and M. Shah, "Learning motion patterns in crowded scenes using motion flow field," in 19th Int. Conf. Pattern Recognition, Tampa, FL, Dec. 8–11, 2008, pp. 1-5.

[20] R. Mehran, B. E. Moore, and M. Shah, "A streakline representation of flow in crowded scenes," in *Proc. 11th European Conf. Computer Vision*, Heraklion, Greece, Sep. 5–11, 2010, part III, pp. 439–452.

[21] W. Wang, W. Lin, Y. Chen, J. Wu, J. Wang, and B. Sheng, "Finding coherent motions and semantic regions in crowd scenes: A diffusion and clustering approach," in *Proc. 13th European Conf. Computer Vision*, Zurich, Switzerland, Sep. 6–12, 2014, part I, pp. 756–771.

[22] A. B. Chan and N. Vasconcelos, "Modeling, clustering, and segmenting video with mixtures of dynamic textures," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 5, pp. 909–926, 2008.

[23] D. Kuettel, M. Breitenstein, L. Van Gool and V. Ferrari, "What's going on? Discovering spatio-temporal dependencies in dynamic scenes," in *IEEE Conf. Computer Vision and Pattern Recognition*, San Francisco, CA, Jun. 13–18, 2010, pp. 1951–1958.

[24] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in *IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, Miami, FL, Jun. 20–25, 2009, pp. 935–942.

[25] W. Choi and S. Savarese, "A unified framework for multi-target tracking and collective activity recognition," in *Proc. 12th European Conf. Computer Vision*, Florence, Italy, Oct. 7–13, 2012, part IV pp. 215–230.

[26] C. C. Loy, T. Xiang and S. Gong, "Detecting and discriminating behavioural anomalies," *Pattern Recognit.*, vol. 44, no. 1, pp. 117–132, 2011.

[27] I. Couzin, "Collective minds," *Nature*, vol. 445, no. 7129, pp. 715–715, 2007.

[28] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, "Novel type of phase transition in a system of self-driven particles," *Phys. Rev. Lett.*, vol. 75, no. 6, pp. 1226, 1995.

[29] M.-C. Chang, N. Krahnstoever, and W. Ge, "Probabilistic group-level motion analysis and scenario recognition," in IEEE Int. Conf. Computer Vision, Barcelona, Spain, Nov. 6–13, 2011, pp. 747–754.

[30] C. Tomasi and T. Kanade, "Detection and tracking of point features," Carnegie Mellon Univ., Tech. Rep. CMU-CS-91-132, Apr. 1991.

[31] B. Schölkopf, J. Platt, and T. Hofmann, *A Local Learning Approach for Clustering*, Cambridge: MIT Press, 2007 pp. 1529–1536.

[32] C. C. Aggarwal, "A human-computer interactive method for projected clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 16, no. 4, pp. 448–460, 2004.

[33] W. M. Rand, "Objective criteria for the evaluation of clustering methods," *J. Am. Stat. Assoc.*, vol. 66, no. 336, pp. 846-850, 1971.

Towards An Effective and Efficient Approximation Algorithm for Advanced Computer Vision Applications based on Two-Dimensional Dynamic Programming

Alfredo Cuzzocrea
DIA Dept., University of Trieste and ICAR-CNR, Italy
alfredo.cuzzocrea@dia.units.it

Enzo Mumolo
DIA Dept., University of Trieste, Italy
mumolo@units.it

Giorgio Mario Grasso
CSECS Dept., University of Messina
Italy
gmgrasso@unime.it

Gianni Vercelli
DIBRIS Dept., University of Genova, Italy
gianni.vercelli@unige.it

Abstract

The Dynamic Programming Algorithm (DPA) was developed in the fifties. However, it is sometimes still used nowadays in various fields because it can easily find the global optimum in certain optimization problems. DPA has been applied to problems of one, two or three dimensions. When the dimension of the problem solved by DPA is equal to one the complexity of the algorithm is polynomial but if the size is greater than one, the complexity becomes NP complete. In such cases a practical implementation of the algorithm is possible only using some approximation. In this paper we present a novel approximation of the two-dimensional Dynamic Programming Algorithm (2D-DPA) with polynomial complexity. We then describe a parallel implementation of the algorithm on a recent Graphics Processing Unit (GPU).

1 Introduction

In this paper we describe an approximation of the two-dimensional Dynamic Programming algorithm (2D-DPA) and its implementation on a GPU device with CUDA architecture. It is well known that the the two-dimensional dynamic programming algorithm is NP-complete. Hence, approximate versions are required for its execution. The approximation we propose here has a polynomial complexity.

The dynamic programming algorithm (DP) is based on the principle of optimality of Richard Bellman [3] and it is an elegant method for finding the global solution to certain optimization problems. Although the optimization problem with dynamic programming was originally formulated as a continuous variational problem, later it was first discretized and then solved as a combinatorial problem (namely discrete) optimization [1]. The application of the dynamic programming algorithm requires that an optimization problem is formulated as a series of simpler problems. The global optimum is obtained by the sequence of local optima. A classic example is to align two sequences of symbols, which has found application in speech recognition [21] and bioinformatics [30].

DP has been applied to various problems of pattern recognition and computer vision [1, 9]. Although there are many other optimization techniques, DP is an ideal technique to solve many discrete optimization problems such as inventory management, scheduling of activities or packaging of objects. Recently Buchanan and Fitzgibbon [4] describe an algorithm that processes multiple hypotheses for tracking mobile objects with dynamic programming. In [31, 15] methods for computing the disparity map of stereo images with dynamic programming are described. Furthermore, the elastic comparison between images as described for example by Uchida [29] is a typical application of the two-dimensional dynamic programming.

While DPA was originally used as a method to efficiently solve optimization problems, [3], Angel [2] uses the analytical DP to smooth iterpolated data. Serra and Berthod [22] and Munich and Perona [19] use the DP algorithm for aligning non-linear one-dimensional patterns. Most recently, Uchida *et al.* [25] use DPA for tracking objects.

The algorithm of DP (and its extension, the stochastic programming, namely the Hidden Markov Models) is a classic technique for recognizing the spoken voice [21] and for the recognition of printed characters [17].

DOI reference number: 10.18293/DMS2016-031

Many researchers have extended one-dimensional dynamic programming algorithms (1D-DP) to the two-dimensional case. The comparison of images by two-dimensional DPA has been described for example in [16, 26], but the authors have met the computational difficulties due to the NP completeness of the problem [14].

To address this computational difficulty, different approximation strategies have been proposed, with the aim of arriving at a solution with polynomial complexity but with a lower level of optimization. The approximation type described in [13] limits the flexibility of the correspondences. Another approximation strategy consists in the partial omission of the mutual dependence between four adjacent pixels, such as the tree representation described in [18].

All elastic comparison algorithms between patterns are based on dynamic programming as a method of combinatorial optimization. A recent overview by Felzenszwalb *et al.* [10] emphasizes that the use of dynamic programming as a discrete optimization method is versatile, and arises in very different low-level and high-level vision problems.

This paper is organized as follows. In Section 2 various applications based on DPA implemented on CUDA architecture are described. Section 3 describes dynamic programming algorithms in one and in a two dimensions. In Section 4 we describe a two-dimensional DPA approximation algorithm with polynomial complexity. This as an improvement of well-known argumentation that state that 2D-DPA has an exponential complexity. Section 5 reports the CUDA-based implementation of our proposed approximate 2D-DPA algorithm. Finally, in Section 6, we report concluding remarks and hints on future work. An extended version of this work appears in [7].

2 Related Work

The two-dimensional dynamic programming algorithm requires a large number of computations. For this reason many authors have implemented the algorithm in parallel form on Graphics Processing Devices (GPU). The main problem that have been considered is how to find the best way to parallelize the 2D-DPA.

Many problems have been solved with the 2D-DPA algorithm. For example, the problem of finding the disparity between stereo images, the problem to compute a distance between images using the so-called elastic matching method, or several discrete numerical problems have been addressed using the two-dimensional dynamic programming algorithm. In 2007, a dynamic programming algorithm to solve the problem of finding the disparity between stereo images has been implemented on an ATI Radeon X800 GPU, one of the first GPU devices [11]. In 2009, Xiao *et al.* [32] address the problem of how to map the dynamic programming on a graphics processing unit. They propose a fine-grained parallelization of a single instance of the DP algorithm that is mapped to the GPU. In the same year Congote *et al.* [6] compute the map of depth from stereo images for three-dimensional display on a GPU using dynamic programming. They use an NVIDIA GPU commercially available in that year, in particular the model GTX 295, achieving an execution speed of 25 frames per second. The heart of their work was the use of optimized algorithms for a GPU execution with stable depth maps and little noise. In 2010, Steffen *et al.*, [23], describe their change to the numerical software environment called Algebraic Dynamic Programming consisting of different optimization problems solved with dynamic programming. Their change has been the implementation of the dynamic programming based algorithms on a GTX 280 GPU. The authors report that the speed-ups range, depending on the application, from about 6 to about 25. Stivala *et al.* [24] published an article in 2010 that shows how to parallelize any DPA on a multicore computer with shared memory by means of a hash table without using primitives for mutual exclusion lock. The authors adopted this strategy: initially they create multiple threads that compute top-down DP recursion storing the result in a hash table with no spin-lock.

3 One- and Two-Dimensional DPA

Dynamic Programming is often explained using the *edit* distance, which measures the number of insertions, deletions and substitutions required for matching two sequences of symbols [20]. In other words the edit distance is a way to measure the similarity of two strings or to align two strings. The *edit* distance has many applications, for example in bioinformatics [8] or in natural language [12]. The *edit* distance algorithm is described as follows. Given two one-dimensional sequences, $A = (a_1, a_2, ..., a_i, ..., a_N)$ and $B = (b_1, b_2, ..., ..., b_j ... b_M)$, the path M' from the cell $(1, 1)$ to the cell (N, M) of the $A - B$ space that gives the minimum accumulated distance represents the mapping of one sequence relative to the other. The path is formed by a series of points in such a way that each point k of the path corresponds to a pair of coordinates of the two sequences, $M'_k = (i_k, j_k)$. The distance between the two sequences is the sum of the local distances $d(M'_k)$ between the two addressed elements of the sequences, a_{i_k}, b_{j_k}, computed along the path M', or $\sum_{k=1}^{|M'|} \|a_{i_k} - b_{j_k}\|$, where $|M'|$ is the length of the path M', namely the number of points k. The distance $\|.\|$ depends on the problem under examination.

The mapping between the two sequences is the path along which the final accumulated distance is minimum. Finally, the distance between the two sequences is the mini-

mum accumulated distance:

$$D(A,B) = \frac{\min_{M'} \sum_{k=1}^{|M'|} d(M'_k)}{|M'|} =$$

$$= \frac{\min_{M'} \sum_{k=1}^{|M'|} d(i_k, j_k)}{|M'|} = \frac{\min_{M'} \sum_{k=1}^{|M'|} \|a_{i_k} - b_{j_k}\|}{|M'|} \quad (1)$$

We note that the factor at the denominator is used to normalize the distance in respect of the length of the optimal path, which may be different. The optimization problem described in (1) is solved by the dynamic programming algorithm. In each point of the $A - B$ space the accumulated cost distance $D(i,j)$ is updated using the recursion shown in (2).

$$D(i,j) = min \begin{cases} D(i-1,j-2) + 2d(i,j-1) + d(i,j) \\ D(i-1,j-1) + 2d(i,j) \\ D(i-2,j-1) + 2d(i-1,j) + d(i,j) \end{cases} \quad (2)$$

where $D(0,0) = 2d(0,0)$. The recursion (2) expresses the optimality principle of dynamic programming. This type of recursion is the *symmetric* form of dynamic programming. It is important to remark that using the recursion (2) the length of the optimal path length just mentioned above, which is the normalization factor, is the sum of the lengths of the sequences A and B, namely $N + M$. The recursion reported in (2) is graphically shown in Fig. 1.

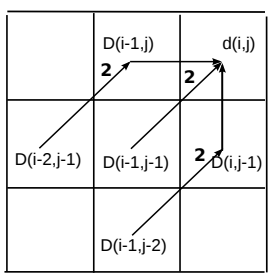


Figure 1. Graphical Representation of DP Recursion

The recursion described in (2) can be implemented according to the following pseudocode, where we call U, V, W, the accumulated cost distances computed at the points, respectively, $D(i-1,j-1)$, $D(i,j-1)$ and $D(i-1,j)$.

```
for i = 0 to N
  for j = 0 to M
  {
    Wval=score[j]; Wfree=free[j];
    if( (Wval+dij)<(Uval+2dij)) && Wfree)
      if(Wval<=Vval)||(!free) {
        Vval=Wval+dij;/* W */
        Vfree=0;
      }
      else {
        Vval=Vval+dij; /* V */
        Vfree=0;
      }
      else {
        if(Vval+dij)<(Uval+2dij) && Vfree)
        {
          Vval = Vval + dij;/* V */
          Vfree = 0;
        }
        else {
          Vval=Uval+2dij;/* U */
          Vfree = 1;
        }
        score[j]=Uval; free[j]=Ufree;
        Uval=val;
  }
  final distance = score[M]/(N + M);
```

As shown in the above pseudocode, after the recursion (2) is applied to all the (i,j) points of the $A - B$ space, equation (1) becomes:

$$D(A,B) = \frac{D(N,M)}{N+M} \quad (3)$$

where $D(N,M)$ is the final distance between the two one-dimensional sequences, namely the distance accumulated to the point (N,M). The sum $N + M$ at the denominator is the normalization factor.

It is important to remark that the constraints shown in Fig.1 and represented by the *Ufree*, *Vfree* and *Wfree* variables of the above pseudocode cause that the optimum mapping path can not have two consecutive horizontal or vertical movements. Thus, as stated in [21], unrealistic mapping paths are avoided.

Since the path M' is the optimal mapping between the two sequences, it can be used to align one sequence onto the other, which can be performed by inserting or removing a point of one sequence according to the horizontal or vertical moves. This operation is called warping. The goal of the warping is to stretch or shrink one sequence to make it identical to the other.

We now extend the derivation of a distance between two one-dimensional sequences given above to the derivation of a distance between two two-dimensional sequences, namely the images, $X = \{x(i,j)\}$ and $Y = \{y(u,v)\}$. For that, we introduce a mapping plane M", $M" = \{m(k,l)\}$, which maps the two images from the first row of the two images to the last row of the two images. Each of the elements of M" correspond to a pair of coordinates of pixels of the two images, namely $M"_{k,l} = m(k,l) = ((i,j)_{k,l}, (u,v)_{k,l})$. Similarly to the one-dimensional case, a distance between the two images can be defined as reported in (4).

$$D(X,Y) = \frac{\min_{M"} \sum_k \sum_l d(M"_{k,l})}{|M''|} =$$

$$= \frac{\min_{M"} \sum_k \sum_l \|x(i,j)_{k,l} - y(u,v)_{k,l}\|}{|M"|} \quad (4)$$

where, as before, $|M"|$ is a normalization factor.

The plan $M"$ that solves equation (4) is a mapping between the two images. Using the mapping $M"$ one can stretch and shrink an image for overlapping it onto the other one. Similarly to the one-dimensional case this operation is called warping between images.

However, it has been shown that the optimization described in (4) is NP-complete. As previously mentioned, many authors have developed 2D-DPA algorithms using various approximation strategies to make them computationally tractable. The algorithm developed by Levin and Pieraccini in 1992 has a complexity of $O(N^{4N})$, assuming that N is the height and width of the images [16]. Uchida *et.al* described in [27] and in [28] a two-dimensional dynamic programming algorithm with $O(N^3 9^N)$ complexity. Uchida and Sakoe described in [29] various elastic matching algorithms proposed so far, seven of which are based on dynamic programming.

In the following section we describe an algorithm with a complexity of $O(N^4)$.

4 Approximate Two-Dimensional DPA

In this Section, we provide the main contribution of our research, i.e. the approximate 2D-DPA.

The algorithm proposed here for the mapping of images is based on the one-dimensional DPA described in Section 3. Consider an image as a vector whose elements are the rows of pixels of the image itself. Let us indicate with $x(i,:)$, $y(i,:)$ the i-th row of pixels of the images X, Y. The X, Y images are thus described as reported in (5).

$$\begin{aligned} X &= [x(1,:), x(2,:), ..., x(i,:), ..., x(N,:)]^T \\ Y &= [y(1,:), y(2,:), ..., y(j,:), ..., y(N,:)]^T \end{aligned} \tag{5}$$

In (5) the images are assumed for simplicity of the same size. The idea of this paper is to apply the one-dimensional DPA algorithm on the two sequences X and Y. We remark that each element of these sequences is an entire row of pixels. The $i-$ th row of X is $x(i,:) = (x_{i,1}, \ldots, x_{i,n}, \ldots, x_{i,N})$ and the $j-$ th row of Y is $y(j,:) = (y_{j,1}, \ldots, y_{j,m}, \ldots, y_{j,N})$. The distance between two elements of X, Y or, in other terms, the distance between two rows of pixels is again performed with one-dimensional DPA. The application of (1) to $x(i,:)$, $y(j,:)$ becomes (6).

$$\begin{aligned} d(x(i,:), y(j,:)) &= \frac{\min\limits_{M'} \sum_{l=1}^{|M'|} d(M'_l)}{|M'|} = \\ &= \frac{\min\limits_{M'} \sum_{l=1}^{2N} \|x_{i,n_l} - y_{j,m_l}\|}{2N} \end{aligned} \tag{6}$$

On the other hand, the application of 1 to X, Y results in (7). In this case the map $\overline{M'}$ is between all the rows of X and Y. As before, $|\overline{M'}|$ is the length of the path of the $\overline{M'}$ map.

$$D(X,Y) = \frac{\min\limits_{\overline{M'}} \sum_k d(\overline{M'}_k)}{|\overline{M'}|} = \frac{\min\limits_{\overline{M'}} \sum_k d(x(i,:), y(j,:))}{|\overline{M'}|} \tag{7}$$

Substituting (6) in (7) we obtain the final expression reported in (8).

$$\begin{aligned} D(X,Y) &= \frac{\min\limits_{\overline{M'}} \sum_k \frac{\min\limits_{M'} \sum_l d(M'_l)}{2N}}{2N} = \\ &= \frac{\min\limits_{\overline{M'}} \{\sum_{k=1}^{2N} \min\limits_{M'} \sum_{l=1}^{2N} \|x_{i,n_l} - y_{j,m_l}\|}{4N^2} \end{aligned} \tag{8}$$

Clearly, the two min operators represent the fact that for computing the optimum path between images with DPA, other optimum paths between the rows are computed with DPA. Recall that we assumed the images are $N \times N$ pixels, then the length of the optimal path between the two images is $2N$. Local distances at any point in this process are obtained with other 1D-DPA with paths length equal to $2N$. The total length is the sum of $2N$ along the long path $2N$, giving the term $4N^2$ at the denominator of (8).

The algorithm is approximated since it is not guaranteed that the warping function is continuous. In fact, the one-dimensional DPA aligns the rows independently of each other. This means that the algorithm can be successfully applied if the difference between the two images is somehow limited.

5 CUDA-based Implementation of Approximate 2D-DPA

Our approximate realization of 2D-DPA is the following: we compute a 1D-DPA between each row of an image and each row of the other image. A matrix of distances between rows is then computed. On this matrix another 1D-DPA is applied. This process is described by equation (7). In Fig. 2 we show how the distance matrix is computed: between i-th row of the first image and the j-th row of the second image a 1D-DPA is computed to fill the (i,j) matrix element.

Let us look at Fig.3. It represents a matrix of threads. The elements $x_{i,0}, \ldots x_{i,3}$ is the $i-$th row of the X image, assuming that its length is 4, and the elements $y_{i,0}, \ldots, y_{i,3}$ is the $j-$th row of the Y image.

Each cell of the matrix computes the DPA recursion. Thus, the accumulated distance at cell $(3,3)$ is the distance between the two rows. In Fig.3 we indicate with arrows the

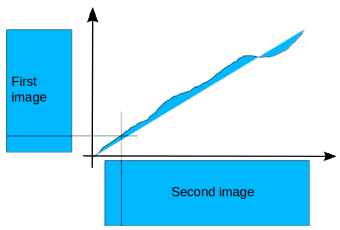


Figure 2. Approximate Mapping Between Two Images

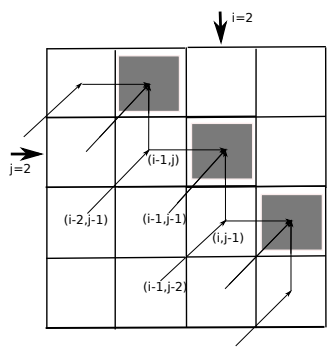


Figure 3. A Matrix of Threads

element (2,2). The corresponding element is indicated with a black square. it is obvious from the DP recursion that the computation of the accumulated distance $D(2,2)$ can be performed only if the distances in the elements whose coordinates are marked between round brackets are already available. This means that the only possible sequence of elements that can be computed in parallel, i.e. simultaneously, is the sequence of elements drawn in black in Fig.3. The parallel computation of elements to the final (M,N) point thus proceeds according to diagonal sequences of elements as depicted in this figure.

Fig. 4, shows how the algorithm can be mapped on the GPU. In this figure, on the vertical axis, a series of image rows, each with 4 elements, are reported. The horizontal models are processed one at a time, from left to right. The figure shows the situation relating to the third horizontal element. The figure shows that all cells drawn black in the Fig. 4 can be computed simultaneously.

From the above it is evident that the parallel implementation is implemented taking account of all models at once. This is accomplished with a single `while` cycle exploring the number of iterations required to complete the matrix minimum, whose size is equal to `DimX` + `dimy- 1`, where `DimX` and `dimy` are the width and height of the matrix.

Since now all of the comparisons matrices are considered simultaneously, the number of iterations needed to complete the calculation is the size of the larger matrix.

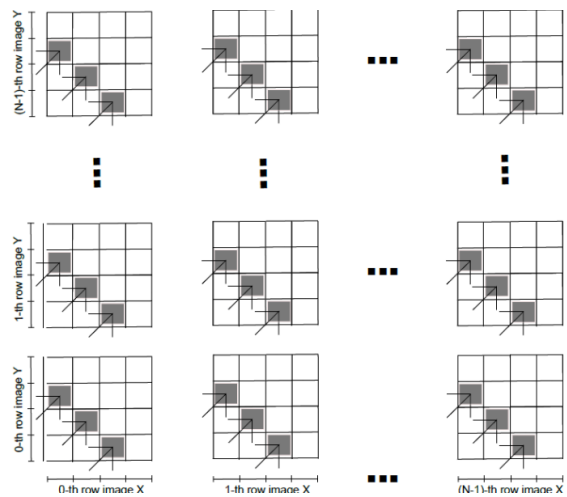


Figure 4. Simplified Representation of the Mapping af the Algorithm on GPU – each black cell represents a thread; horizontal and vertical patterns will be image rows

The following pseudocode shows how we call the kernel.

```
for (r=0;r<N;r++) {
    for (iter=0;iter<=N;iter++) {
        1D-DPA<<N, iter+1>>>(r,iter, d_cell, ...);
    }
    for (iter=N-1;iter>=1;iter--) {
        1D-DPA<<<N,iter>>>(r,iter, d_cell, ...);
    }
}
```

The rows of an image are indicated by the variable r, which is reported in the outer loop. The one-dimensional dynamic programming used to compute the distance from one row to the other is carried out in the inner loops. The first inner loop creates a thread for each block, the second creates two threads, the third creates three threads and so on, up to N.

At this point it is important to point out three things.

- first, to prevent the transfers of images from the Host, which would slow down the calculation, all the images are initially loaded in the global memory of the GPU
- second, the operation realized by the above pseudo code computes the matrix of local distances between all the rows of the two images. These distances are used in a final dynamic programming to calculate the distance between the two images as shown by (7).
- the third observation is that if we want to perform the warping between the two images, in addition to the

computation of the distance between the two images, we have to recover the mapping plane between the two images. This is achieved by backtracing last initial cell. This means that when calculating the dynamic programming all the information about the algorithm choices must be stored in appropriate data structures. These data structures are used during the backtracing.

In fact, the matrix `d_cell` is a super-matrix that contains all the matrices of the individual comparisons. The matrix `d _cell` is a three dimensional matrix with indices `x`, `y`, `z` necessary to specify how to make comparisons. The main difference in the code of the function `1D-DPA` is the mapping of the block indexes in the three-dimensional coordinates of the matrix. Of course in addition to storing the decisions taken by the algorithm when processing the distance matrix we must also store those taken during the confrontation between the lines.

In the following we report the pseudocode of the kernel that runs the `1D-DPA` algorithm between the lines of the two matrices on the GPU.

```
__global__ void 1D-DPA(int nrtemp, float *d_pdati, float
      *d_cell, ...)
{
  int z=blockIdx.x/N;
  int y=blockIdx.x \% N;
  int x=d_x[blockIdx.x];
  int i=y + x*N + (d_FrCum[z]*N);
  float dxy, Uval, Vval, Wval;
  char Wfree, Vfree;
  if ((d_checkflag[i]==1) || (y==0 && x==0)) {
      if ((y==0) && (x==0)) {Uval=0; Vval=1000; Wval
           =1000; Wfree=0; Vfree=0; }
      else if ((y==0) && (x!=0)) {
           Uval=1000;
           Vval=1000;
           Wval=d_cell[i-N];
           Wfree=d_bWfree[i-N];
           Vfree=0;
      }
      else if ((y!=0) && (x==0)) {
           Uval=1000;
           Vval=d_cell[i-1];
           Wval=1000;
           Wfree=0;
           Vfree=d_bWfree[i-1];
      }
      else {
           Uval=d_cell[i-N-1];
           Vval=d_cell[i-1];
           Wval=d_cell[i-N];
           Wfree=d_bWfree[i-N];
           Vfree=d_bWfree[i-1];
      }
      dxy=d(z,nrtemp,x,y,d_pdati);// distance between
           pixels
      if ( ( (Wval+dxy) < (Uval+2*dxy) ) && ( Wfree==1 )
            )
           if ( ( Wval <= Vval ) || ( bfree==0) ) {
             Vval=Wval+dxy; // choose W
             d_bWfree[i]=0;
           }
           else {
             Vval=Vval+dxy; // choose V
             d_bWfree[i]=0;
           }
           else {
             if( ( (Vval+dxy) < (Uval+2*dxy) ) && (
                  Vfree==1 ) ){
                  Vval=Vval+dxy; // choose V
                  d_bWfree[i]=0;
             }
             else {
                  Vval=Uval+2*dxy; // choose U
                  d_bWfree[i]=1;
             }
           }
      d_cell[i]=Vval;
      if (x<d_nftemp[z]-1) {
         d_checkflag[i+1]=1;
         d_checkflag[i+N]=1;
         d_x[blockIdx.x]=d_x[blockIdx.x]+1;
      }
      else if (x==d_nftemp[z]-1 && y<N-1) {
         d_checkflag[i+1]=1;

      }
      else if (x==d_nftemp[z]-1 && y==N-1) {
         d_res[z]=Vval;

      }
  }
}
```

There are other important points to consider.

- Recall that when a kernel is started, all the threads execute the same code. However not all of the threads can execute simultaneously, but only those for which the input data have been already computed. So we have to solve this kind of synchronization. The solution that was adopted in this work is simply to use Boolean variables, in particular the variables $d_c heck flag[]$, which are set equal to true when the data are available, and false otherwise.

- The second observation is related to the use of the flag $Wfree$ and $Vfree$. These flags are used to impose a path that can not have two consecutive horizontal or vertical movements. These paths are in fact considered incorrect.

6 Conclusions and Future Work

In this paper we describe an approximation of the two-dimensional dynamic programming algorithm in order to make it computationally feasible. The algorithm has been mapped on a recent GPU device. A first example of application of the algorithm can be the automatic analysis of the inclination of handwritten characters. Finding the degree of inclination could be used to infer the physical or mental state of the writer. This work will continue to develop other types of mapping the 2D-DPA on the CUDA architectures in order to reduce the execution time. On the other hand, another research direction consists in enriching the proposed framework with innovative features such as *adaptiveness* (e.g., [5]) and support for *big data processing* (e.g., [33]).

References

[1] A. A. Amini, T. E. Weymouth, and R. C. Jain. Using dynamic programming for solving variational problems in vision. *PAMI*, 12(9), 1990.

[2] E. Angel. Dynamic programming for noncausal problems. *IEEE Trans. on Automatic Control*, 1981.

[3] R. Bellman. *Bynamic Programming*. Princeton University Press, 1957.

[4] A. Buchanan and A. W. Fitzgibbon. Interactive feature tracking using K-D trees and dynamic programming. In *2006 IEEE (CVPR 2006), 17-22 June 2006, New York, NY, USA*, pages 626–633, 2006.

[5] M. Cannataro, A. Cuzzocrea, and A. Pugliese. XAHM: an adaptive hypermedia model based on XML. In *Proceedings of ACM SEKE 2002, Ischia, Italy, July 15-19, 2002*, pages 627–634, 2002.

[6] J. Congote, J. Barandiarán, I. Barandiaran, and O. E. Ruiz. Realtime dense stereo matching with dynamic programming in CUDA. In *XIX Spanish Computer Graphics Conference, CEIG 2009*, pages 231–234, 2009.

[7] A. Cuzzocrea, E. Mumolo, D. Pirrò, and G. Vercelli. An efficient cuda-based approximate two-dimensional dynamic programming algorithm for advanced computer vision applications. In *IEEE SMC 2016, Budapest, Hungary, October 9-12, 2016*, 2016.

[8] C. Di Neil and P. Pevzner. *An Introduction to Bioinformatics Algorithms*. MIT Press, 2004.

[9] P. F. Felzenszwalb and R. Zabih. Dynamic programming and graph algorithms in computer vision. *PAMI*, 33(4), 2011.

[10] P. F. Felzenszwalb and R. Zabih. Dynamic programming and graph algorithms in computer vision. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(4):721–740, 2011.

[11] M. Gong and Y.-H. Yang. Real-time stereo matching using orthogonal reliability-based dynamic programming. *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 16(3):879–884, 2007.

[12] C. J. Hopfe, Y. Rezgui, E. Métais, A. D. Preece, and H. Li, editors. *15th International Conference NLDB 2010, Cardiff, UK, June 23-25, 2010. Proceedings*, LNCS 6177. Springer, 2010.

[13] D. Keysers, T. Deselaers, C. Gollan, and H. Ney. Deformation models for image recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(8):1422–1435, 2007.

[14] D. Keysers and W. Unger. Elastic image matching is np-complete. *Pattern Recognition Letters*, 24(1-3):445–453, 2003.

[15] C. Lei, J. M. Selzer, and Y. Yang. Region-tree based stereo using dynamic programming optimization. In *2006 IEEE (CVPR 2006), 17-22 June 2006, New York, NY, USA*, pages 2378–2385, 2006.

[16] E. Levin and R. Pieraccini. Dynamic planar warping for optical character recognition. In *Proceeding of ICASSP*, pages 149–152, 1992.

[17] C. Liu, S. Jäger, and M. Nakagawa. Online recognition of chinese characters: The state-of-the-art. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(2):198–213, 2004.

[18] V. Mottl, S. Dvoenko, and A. Kopylov. Pattern recognition in interrelated data: The problem, fundamental assumptions, recognition algorithms. In *17th International Conference ICPR 2004, Cambridge, UK, August 23-26, 2004.*, pages 188–191, 2004.

[19] M. E. Munich and P. Perona. Continuous dynamic time warping for translation-invariant curve alignment with applications to signature verification. In *ICCV*, pages 108–115, 1999.

[20] G. Navarro. A guided tour to approximate string matching. *ACM Comput. Surv.*, 33(1):31–88, 2001.

[21] H. Sakoe and S. Chiba. *Readings in speech recognition*, chapter Dynamic programming algorithm optimization for spoken word recognition, pages 159–165. Morgan Kaufmann Publishers Inc., 1990.

[22] B. Serra and M. Berthod. Subpixel contour matching using continuous dynamic programming. In *CVPR 1994, 21-23 June, 1994, Seattle, WA, USA*, pages 202–207, 1994.

[23] P. Steffen, R. Giegerich, and M. Giraud. *GPU Parallelization of Algebraic Dynamic Programming*, pages 290–299. Springer Berlin Heidelberg, 2010.

[24] A. Stivala, P. J. Stuckey, M. G. de la Banda, M. V. Hermenegildo, and A. Wirth. Lock-free parallel dynamic programming. *J. Parallel Distrib. Comput.*, 70(8):839–848, 2010.

[25] S. Uchida, I. Fujimura, H. Kawano, and Y. Feng. Analytical dynamic programming tracker. In *Computer Vision - ACCV 2010, Queenstown, New Zealand, November 8-12, 2010*, pages 296–309, 2010.

[26] S. Uchida and H. Sakoe. A monotonic and continuous two-dimensional warping based on dynamic programming. In *ICPR 1998, Brisbane, Australia, 16-20 August, 1998*, pages 521–524, 1998.

[27] S. Uchida and H. Sakoe. A monotonic and continuous two-dimensional warping based on dynamic programming. In *Proc. 14th ICPR*, pages 521–524, 1998.

[28] S. Uchida and H. Sakoe. An efficient two-dimensional warping algorithm. *IEICE Trans. Inf. and Syst.*, 1999.

[29] S. Uchida and H. Sakoe. Survey of elastic matching techniques for hanwritten character recognition. *IEICE Transactions Inf. and Sist.*, pages 1781–1790, 2005.

[30] A. Vajdi, N. Haspel, and H. Banaee. A new DP algorithm for comparing gene expression data using geometric similarity. In *IEEE BIBM 2015, Washington, DC, USA, November 9-12, 2015*, pages 1157–1161, 2015.

[31] O. Veksler. Stereo correspondence by dynamic programming on a tree. In *IEEE (CVPR 2005), 20-26 June 2005, San Diego, CA, USA*, pages 384–390, 2005.

[32] S. Xiao, A. M. Aji, and W. Feng. On the robust mapping of dynamic programming onto a graphics processing unit. In *IEEE ICPADS 2009, Shenzhen, China, December 8-11, 2009*, pages 26–33, 2009.

[33] B. Yu, A. Cuzzocrea, D. H. Jeong, and S. Maydebura. On managing very large sensor-network data using bigtable. In *IEEE/ACM CCGrid 2012, Ottawa, Canada, May 13-16, 2012*, pages 918–922, 2012.

An Introduction to Geographic Rule Semantics

Robert Laurini
Knowledge Systems Institute, USA
and LIRIS, INSA-Lyon,
University of Lyon, France
Robert.Laurini@insa-lyon.fr

Sylvie Servigne
LIRIS,
INSA-Lyon,
University of Lyon, France
Sylvie.Servigne@insa-lyon.fr

Franck Favetta
LIRIS,
Claude Bernard University of Lyon,
University of Lyon, France
Frank.favetta@liris.cnrs.fr

***Abstract* — As business rules are very common in business information systems, it is fundamental to analyze geographic rules in order to include them into geographic knowledge systems. The goal of this paper is, starting from examples taken overall from urban and environmental planning, to examine geographic rule semantics, and in particular to model space by combining logic and computational geometry. They can come from laws, physical laws, socio-economic considerations, from best practices, or from data mining. By defining new concepts such as metarules, superseded rules and jurisdictions, a conceptual model is proposed.**

Keywords-component: geographic information systems, geographic knowledge systems, geographic knowledge, geographic rules.

I. INTRODUCTION

According to Graham [7] and Morgan [11], rules (business rules) should be considered as first-class citizens in computer science. In enterprises, the 'craft' of expert know-how is capitalized in an information system in the form of "business" rules. These rules can then be explained and implemented in applications such as business intelligence in software architectures integrated type ERP (Enterprise Resource Planning) or not. For example, the SAP integrated software package is based on a declarative formalism for the description of the job tasks with lists of rules, such as "automobile insurance does not cover drivers who have been recognized guilty of driving while intoxicated over the past two years", or "when its monthly invoices are sent, should be included the supplementary documents that match the profile of the client". On the other hand, the explanation and the formalization of business rules is still a hot topic as a new standard from the OMG issued in September 2015 entitled Decision Model and Notation. The objective is to provide a language of formalization of business rules including exploited in the decision-making process (OMG - DMN 2015). For more details refer to http://www.omg.org/spec/DMN/

Thus, a rule is a basic element of a strategy to build reasoning. In contrast to algorithms, they are expressed declaratively. Among business rules, Dietz [6] distinguishes between three categories:

- rejectors typically those related to quality control, that allow a rejection (rejection rules),
- producers such as those determining new values (ex VAT calculation); they can be considered as rules of production of information,
- and projectors such as those related to the replenishment of stocks.

In our case, a rule is not necessarily a legal regulation, but only an inference (implication) between elements or phenomena whose origin can also be physical, statistical or best practice type, or from data mining. Mathematically, a rule will be written in the form of Modus Ponens $A \Rightarrow B$.

Few studies have been conducted on the automatic reasoning in geoprocessing. Nevertheless one can include the works presented in the book edited by Kim *et al*. [9] and the special issue of the CEUS (Computers, Environment and Urban Systems) journal including the editorial of Batty [1] which were very typical of the time. But these works are old and the knowledge engineering framework has drastically changed. More recently, mention the paper written by Jain and Payal [8], but it does not address the reasoning part based on declarative rules.

Include also the paper [13] which operates a GIS on minerals data to discover characterization rules that can facilitate data mining and spatial analysis. In the same spirit, Biasotti *et al*. [2] present methods for extracting the characteristics of a field (peaks, passes, edges, etc.) from the raw digital terrain models, which could easily be described in the form of rules, including those based on the value of the Gaussian curvature of terrains.

However in GIS, beyond administrative rules, other statements may lead to geographic rules. In fact, let us look at some of them:

- in the United Kingdom, we drive on the left;
- in Canada, the majority of the population lives along the border with the United States;
- each capital city has an international airport nearby;
- between the two capitals, in general, there are direct flights;
- in the Northern Hemisphere, the more you are going to the north, the colder (but locally this is not always true).
- the more you climb a mountain, the colder;
- heavy rain upstream, downstream flooding.
- mosques are oriented towards Mecca;
- if a zone is a swamp, it is necessary to prohibit construction;
- if there is unemployment, the creation of companies or industrial areas must be encouraged;
- If a plot is adjacent to an airport, it is necessary to limit the height of buildings;
- It is forbidden to open a new pharmacy within 500 meters of an another already existing;
- A good practice in Mexico is to use a bus to go from Puebla to Oaxaca City.

DOI reference number: 10.18293/DMS2016-015

Another example is derived from the use of electrical plugs in the countries. This table could be considered as a Visual decision table (Figure 1).

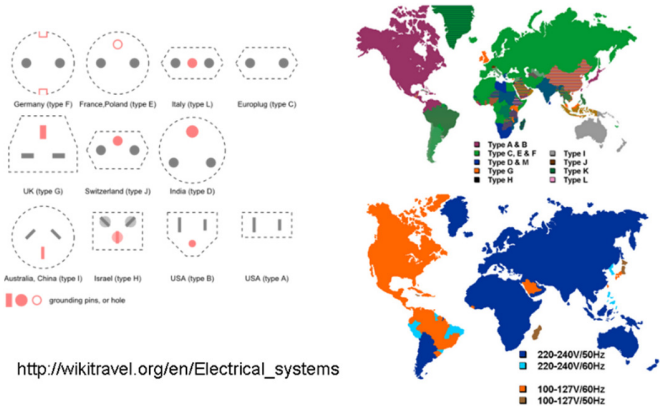


Figure 1. Types of electrical plugs (a) and the countries where they are standardized. Source: http://wikitravel.org/en/Electrical_systems.

Among the rules, there are several categories. Take two examples "if it rains, I get wet" and "if it rains, I take an umbrella". In the first case, it is a rule of physical type whose consequence is systematic provided that I am outside. For the second, it reveals a good practice that I am not obliged to follow. Therefore, we can see that if the premises are identical, the status of the conclusions can be totally different.

Thus, the goal of this paper will to review the rules in geoprocessing in order to extract particular semantics and to lay the foundations of a machine-processable language and model. Voluntarily, we will treat neither temporal aspects nor 3D aspects although sometimes it is necessary to use these dimensions in certain rules.

The areas of applications fall within geoprocessing. Specifically, we are interested in the rules about not only geographic objects, but also about all objects whose knowledge of the location is important.

This paper is thus built: after having presented an introductory example and detailed computer modeling of rules, many of them will be examined. Finally a sketch of model will be given and some perspectives will be open. And to conclude this analysis, a definition of geographic rules will be presented.

II. Introductory Example:Street Naming

In [15], an example of rule encoding is given concerning road naming in Australia in order to automate the process. Rules are defined in the form of ontological vocabularies using SWRL a Semantic Web Rule Language based on a combination of the OWL DL and OWL. For more details, refer to https://www.w3.org/Submission/SWRL/ .

In this paragraph, only some rules are presented. Rule R1 automatically infers information with the help of a road link between proposed and existing roads; this rule is necessary as every road needs to link with at least one other road to allow access. Rule R2 checks road length against road type; checking the road length for shortest road types ('Place', 'Close' and 'Lane') is necessary to avoid confusion with the preference for road usage. Rules R3 and R4 check the compatibility between road usage and road links; for example an open-ended road must have a road link at both start and end points of the road. And finally Rule R5 checks whether or not the proposed road has a wide panoramic view across surrounding areas.

However several remarks can be done because this is logic reasoning, not geographic reasoning:

- Rule 1 tests whether the new proposed road is linked to another existing road; but the link can be by a set of new proposed roads. The adapted solution is concerned by the order of presentation of new streets. A very general solution must be based on graph theory.
- In Rule 2, road length which must be taken into account, is given as a given attribute, not computed from road coordinates.
- In Rule 5, panoramic view is also given as an attribute, not calculated taken terrain morphology into account by 3D computational geometry.

As a consequence from this rapid analysis, classical logic is not sufficient to represent geographic rules since other mathematical domain must be integrated for reasoning.

III. Geographic Knowledge and Reasoning

The objective of this section is to brush the outline between the bases of geographic knowledge and automatic reasoning. After some generalities, we will quickly review the structuring of the knowledge bases and computer modeling of rules.

A General information

The purpose of a declarative rule-based model is to allow automatic reasoning. In contrast to expert systems of the past that were using simply logic, in our case, it will be very different. Here are some examples:

- defining the location of a new airport, a new hospital, a new stadium, social housing, etc. (Computational Geometry and Operation Research),
- checking the compliance of a building vis-à-vis building regulations (Topology and Computational Geometry),
- determining the best way to go from *A* to *B*, (Graph Theory or Computational Geometry in terrains),
- organizing a policy about urban green spaces (Spatial Analysis),
- determining transit policy (Spatial Analysis),
- flood analysis (Differential Equations, Computational Geometry)
- limiting crime in a city (Spatial Analysis),
- building garbage collection circuits (Graph Theory and Spatial Analysis),
- determining the most polluted, noisiest, places (Spatial Analysis and Operation Research),
- organizing the evacuation of the people during a volcanic eruption (Spatial Analysis and Operation Research),
- etc.

From these examples, we find that the geographic reasoning must invoke other mathematical disciplines such as:

- spatial reasoning by the integration of topology and computational geometry,
- graph theory in some cases like search for routes or circuits,
- spatial analysis,
- simulation techniques,
- fuzzy logic and reasoning,
- multi-criteria decision theory,
- and operations research.

These elements can be integrated in the form of procedures which will be invoked at the right time.

B Geographic knowledge bases

Once selected a human language and a territory, remember (Figure 2) that a geographic knowledge base may include an ontology, a gazetteer, geographical objects, the relationship between these objects, rules and mathematical models. Figure 3 shows such a base linked to a geographic inference engine that will allow reasoning.

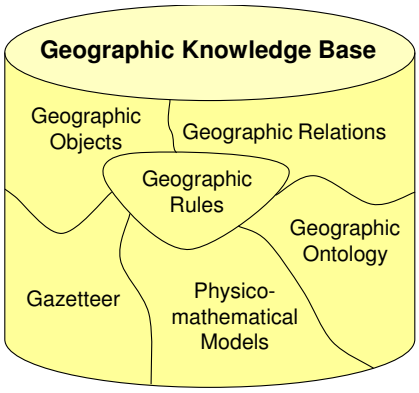


Figure 2. Contents of a geographic knowledge Base.

In input, there will be a territorial project to study so as to determine, through the geographic inference engine, the consequences in output. These consequences will be grouped in a result of feasibility that may be in the form of maps, diagrams, texts, etc.

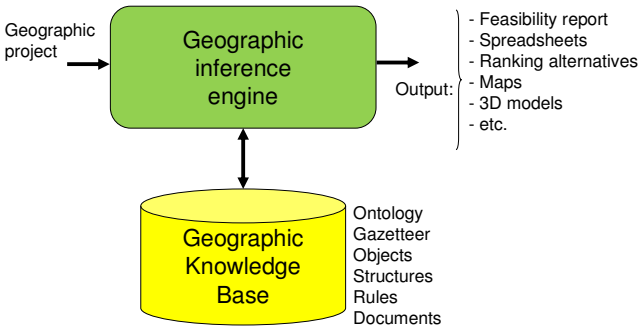


Figure 3. Linking a geographic inference engine to a geographic knowledge base.

To conclude this section, all the components of a geographical knowledge base can be used to write not only a rule but also in their activation context.

IV. Study of the Semantics of Some Geographic Rules

Now let us examine certain rules in different areas. First, we will discuss the rules across the globe (often as physical nature), then those only valid in certain places (often of an administrative nature) and those from data mining. We will continue by rules related to cartographic input and presentation data.

A Global rules

In this category, one may include geodetic and physical geography rules.

A.1 Geodetic rules

The rules of this type are valid all over the world because we consider cardinal points. If *A* is north of *B* and *B* is located north of *C*, then *A* is north of *C*, thus by applying a rule of transitivity. But there is nothing north of the North Pole; so therefore transitivity is limited. A similar rule can be written for the South. But for the East and West, transitivity is partial because of the rotundity of the Earth. If Rome is East of Los Angeles, and if Los Angeles, is East of Beijing, Rome should be to the East of Beijing, but it is actually to the West. The comprehensive rule will be written thus:

```
IF East (A, B) and East (B, C) and Longitude
(A, C) < 180° THEN East (A, C) ELSE West
(A, C).
```

Remember that rarely there are lists of type *North (R, S)* or *East (U, V)* as such information is hidden in the coordinates of the *R*, *S*, *U* and *V*.

If the application of such rules is easy for points, it is not the case for areas. If it is common to hear that the Switzerland is East of France, it remains that, given the geometry of the border, some places of the Switzerland are west of some points of the France. In the case of automatic reasoning, more sophisticated definitions should be developed for areas.

A.2 Rules of physical geography

In this area, the rules should represent natural phenomena and their consequences. For example following tsunamis, volcanic eruptions, storms, heavy rain, we must consider some automatic consequences. But, in addition, we must consider more recently, prevention, protection or mitigation, and effective real-time monitoring systems.

But due to the local topography, some rules may be invalidated. In the northern hemisphere, the more you go north, the colder. A certain scale, this type of rule is valid, but there are places where this reasoning is no longer valid. Therefore, one must distinguish between local and global rules. Here the local rule supersedes the global rule as, for example also when one is dealing with microclimates.

The rules of the spatial distribution of flora and fauna, hydrology, etc. from climatology, meteorology fall into this category. Figure 4 gives an example of vegetation layers in the Alps.

Suppose that we are on the southern slopes (sunside). Two options to write a rule:

```
IF Lichen THEN Altitude > 3000;
```

```
IF Altitude > 3000 THEN Lichen.
```

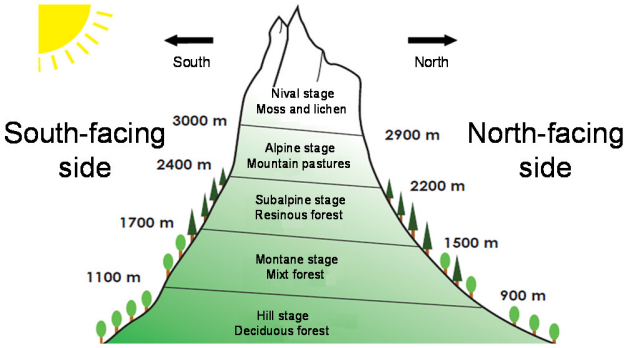


Figure 4. Example of vegetation layers in the Alps. Source: translated from https://www.jardinalpindulautaret.fr/jardin/cadre-naturel-exceptionnel/letagement-vegetation-en-montagne

In the first case, it would be a study linking a type of vegetation to elevation, while the second shows since we are at some level from the sea, what are the types of flora that we can meet.

Now consider the case of mathematical models and assume that we have $A = M\ (B,\ C,\ D)$. This formula can be easily transformed into a rule in the following manner: one must:

- write a procedure or a function representing M, which will be encapsulated,
- search or determine to B, C and D in the antecedent part,
- then run M in the consequent part in order to determine A.

Another variant would be that A may intervene into a condition; therefore M will be invoked in the antecedent part.

B Local rules

Here, we will discuss only the rules applied on a restricted territory, namely the administrative rules and those relating to specific spaces.

B.1 Rules coming from laws

Each country has its own rules, not only from the administrative point of view, but also from location. For instance, when analyzing road traffic from aerial photos, it is important to know that in the United Kingdom the cars drive on the left. In addition, concepts such as language and currency can impact geographic rules.

Typically in addition in each country there are a Constitution and many laws governing the geographic aspects in relation to urban and environmental planning. Take the example of pharmacies in some countries. Figure 5 shows an example of this rule. To establish where it is possible to create a new pharmacy, several geometric operations such as determination of buffer zones and geometric difference have to be implemented in an encapsulated way.

One of the peculiarities of the rules of administrative origin is the existence of sanctions whenever they are not met. But there may be cases of exemptions, exemption which should be taken into account in one way or another.

On the other hand, although there are international standards for the Highway Code, each country has its own peculiarities concerning for instance priority in the intersections, roundabouts, etc.

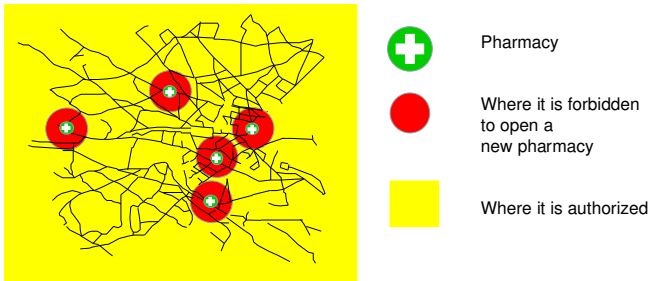


Figure 5. Example of administrative rule: "it is forbidden to open a new pharmacy within 500 meters of another existing one".

B.2 Urban planning rules

Generally, in each country, there are also laws that govern urban planning. Let us take a small example taken from building licenses as presented in Figure 6 where you can see a building that must follow certain rules.

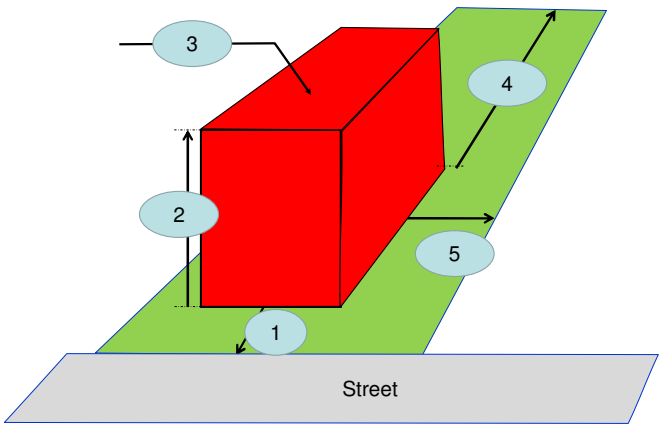


Figure 6. Example of building-related planning rules.

In Figure 6, Rule 1 gives the minimum distance to the road, Rule 2, the maximum height of the building, Rule 3 the volume of the building, Rule 4 the distance to the end of plot and the 5, the distance with neighbors. Here, the rules will be treated as constraints, *i.e.* to be valid (accepted according to the meaning of the regulation), the building project must comply with this set of constraints.

Among the urban planning rules, there may be good practices as "bury the air engineering networks (electricity, telephone)" or even "before to create an underground metro line, we must move sewerage networks".

B.3 Local socio-economic rules

Due to its great inertia, demography generates rules of the type: "the more of children, the more of schools." But may this statement be considered as an administrative rule, a good practice or a recommendation? And what about time frame?

A majority of countries have developed rules for the organization of the economy and companies. These rules have a significant impact on the use of the soil. For instance, consider

the rule "along the edges of sea, the greater the distance from the sea, the lower are prices of homes."

A particular case regards rules related to the flow of people or goods. In these two places or two families of places may intervene, either as origin or destination.

B.4 Good or best practices

Among the rules of good practice, there are those related to the description of itineraries. In most countries where the landscape is charged, description is often done according to the directions, cities and villages to cross. In other countries where the landscape is lightly loaded, the description is made by taking into account numbers and route directions, North, South, East and West. In the deserts and seas, it was common to use the positions of stars as landmarks. These good practices had been used for centuries; but now there are other ways to do so. However this type of knowledge is out of the topic of our research because it uses information outside the terrestrial globe (extra-terrestrial knowledge).

Good practices include techniques of numbering houses in cities, which may vary depending on the country. Alternatively, along a highway, the creation of an additional interchange may be the basis of economic development of a small town.

C Material rules

In this category one can find the rules related to the acquisition and visualization of data. Among these, one can discuss three types of rules:

- those related to data quality control [14, 4, 5],
- and those related to the mutation of object's geometric types (f.i. area to point) and topological relations according to scales (disjoint to meet). For more details, refer to [10, 11]. For instance, the rule mutating an area into a point according to scale (as its centroid), can be written as follows (in which *O* is an area-type geographic object, *2Dmap* a function transforming the object at the scale σ and the ε's some thresholds):

$$\forall O \in OG, \forall \sigma \in Scale, type(O)=area, O_\sigma = 2Dmap(O, \sigma): (\varepsilon_i)^2 < Area\ (O_\sigma) < (\varepsilon_{lp})^2$$

$$\Rightarrow \{Type(O_\sigma)=point;\ O_\sigma = Centroid(O)\}.$$

And then a smaller object can disappear:

$$\forall O \in OG, \forall \sigma \in Scale, type(O)=area, O_\sigma = 2Dmap(O, \sigma): Area\ (O_\sigma) \leq (\varepsilon_{lp})^2$$

$$\Rightarrow O_\sigma = \varnothing.$$

D Rules and plurality of places

Four cases are to be analyzed.

1 / From the previous examples, we can see that most of the rules refer a unique place. But a rule such as "in England we drive on the left" also applies in other countries.

2 / However the rules related to the flow of people, goods and animals are characterized by two places, a so-called origins and destination. Consider for example those governing the movements of migratory birds. Accordingly the grammar of the rules must allow this scenario by considering three cases:

- bipolar flows (the most common),
- diverging flows where only the source is known, for example emigration,
- and the converging flow where only the destination is known, as for immigration.

3 / A third case is that of the rules corresponding to clusters of areas according to certain criteria, such as for example research of homogeneous areas in geomarketing or the definition of electoral boundaries. It is noteworthy that identifiers to these newly-created areas need to be assigned.

4 / And finally another example is about the importation of a good practice from one location to another location.

E Rules and of shareholders' logic

One of the difficulties is the fact that among the urban actors, some have different "logics". With regard to industry creation, an environmentalist or an industrialist may have different ideas on the possible implications of this or that choice. Similarly, some groups may have different priorities: before an empty space, athletes imagine a stadium, pupil's parents a school and a promoter a building, etc.

From the formal point of view, these aspects will occur in multi-actor and multi-criteria decision support systems.

V. TOWARD GEOGRAPHIC RULES MODELING

Now that various examples were analyzed, it is possible to extract elements of modeling. Firstly, general considerations will be given, and then a computer model will be proposed. According to [12], they can be initially modeled by `IF-THEN-fact` or `IF-THEN-Action`. Some `WHEN` or `WHERE` clauses can respectively be applied for temporal and spatial aspects.

As we saw earlier, new concepts have emerged and it is necessary to clarify what is meant by superseded rules, metarules, jurisdiction, etc. Finally two tables will make it possible to synthesize the characteristics of rules and our level of knowledge and their formalization.

A Superseded rules

Indeed, certain rules can be superseded locally. In other words, it will be necessary to take account of this aspect not only in designing the rules, but also in the inference mechanism. This can be written such as: `WHERE SOMELAND IF ISSUE32 THEN REPLACE RULE#25 BY RULE#28.`

B Metarules

A metarule is a rule which arises from other rules for example in a regulatory framework. For example, all local urban plans must be in compliance with some higher level regulations, which thus appear as metarules. In other words, a metarule defines a set of rules that will be valid only when you will refer to this metarule. In addition It can define new concepts, new legal mechanisms, or even of new decision-making bodies; a

metarule can therefore come enriching an ontology with this new terminology; this point although current is extremely complex and will not be treated.

C Jurisdiction

One can call jurisdiction, the territory of application of a rule, a metarule and even the entire knowledge base. Therefore, the gazetteer will only deal with place names within this jurisdiction or through it (rivers, roads, etc.). In some cases, it would be advisable to include close external information such as the names of the neighboring territories.

D Geographic rules and objects

As seen previously, the geographic rules commonly involve geographical objects (e.g. buildings in flood zone) and also geographic objects; see for instance vegetation (Figure 4) or habitats of animals.

But in addition, geographic objects may be deducted from rules. Take the example in maritime laws that distinguish territorial waters and international waters. The principle is based on the distance of 200 nautical miles except for particular cases.

The big problem is that the variables depth and width are not known explicitly, but from a 3D geometric reasoning from the morphology of the waterway. Furthermore, if the river has narrow meanders, it must ensure that long barges can pass.

Ultimately, we deal with producing rules according to Dietz' terminology [6]. Generalizing the previous examples, we need to integrate those items deducted geometrically from geographic rules, namely new objects, new types of objects, new attribute values, or even new relationships between two objects and new spaces that can intervene *e.g.* as jurisdiction.

Thus, once known the rule for the determination of these objects, we can create a new class from this geographic rule. In other words, the consequence part of the rule will enrich the ontology by a creation of new class derived from geometric reasoning, and sometimes even enrich the gazetteer.

E Summary

At the level of formalization, rules of physical geography can be encapsulated into programs from a procedural way we need a mechanism to integrate them into a declarative model. With regard to the laws, they are known at a time t, but may change over time in the form of statutory instruments for which the translation into declarative forms may be difficult or too simplistic.

In addition to the previous features, need to clarify the nature of the implication ($\Rightarrow$). In fact, it has several meanings:

- in the case of physical phenomena, it corresponds to physical laws or causal chains (cause and effect); the implication is therefore automatic, sometimes within a delay;
- However, if the physical law is only known empirically, the kind $a = f(b, c, d,...)$, formula where a is the value of an attribute of a geographic object, then the rule will affect the value calculated with a margin of error;
- Rules of legislative type are human laws; generally if the law is not enforced, sanctions may appear, thus involving an `ELSE` clause in the computer rule.
- In association rules (sometimes referred to as frequent association) from data mining, the semantics of the $\Rightarrow$ sign should be modulated according to the value of confidence related to this association rule;
- Concerning good practice rules, the implication will be judged in desirable manner;
- Finally, if the rule involves fuzzy objects, the semantics of the $\Rightarrow$ sign will be modulated according to the values of fuzzy membership degrees.

Now that the main elements of geographic semantics have been identified, it seems possible to propose a first model.

F Outline of model

From the analyzed examples, first there is that there is a many-to-many relationship between the rules and the names of places, and another between the rules and the types of geographical objects.

In addition, some sites have the ability to emit the metarule (countries) which will apply on inner places and will be a normative framework for located rules. Figure 7 depicts this model. However, regarding rules, things are a little more complicated because they are encoded in a language that remains to be defined, for example from an extension of RuleML [3]. Indeed, for the design of the code of these rules, knowledge must include from not only geographic objects and the relationships between them, but also ontologies (especially for the types of objects), the names of places and mathematical models as shown in Figure 8.

From these observations, to describe a set of geographic rules, in this new language, two levels will be necessary, the rules themselves and their sets. As a first step to simplify the problem, it will not handle metarules and their consequences.

It will be important to include the elements common to all rules included in this set, i.e., the name of the set, language, jurisdiction, ontology and gazetteer. It could include references to other sets of rules provided that the language, ontology and gazetteer are compatible.

Then will be given the rules themselves.

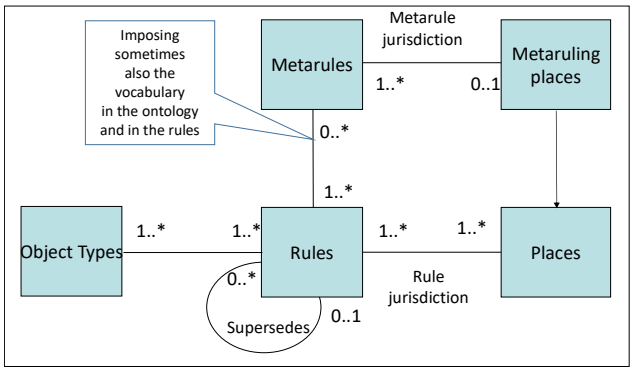


Figure 7. Modeling of geographic rules.

As explained in Figure 8, three parts must be analyzed. In the "antecedent" part, it will be necessary to give the jurisdiction of the rule. This could be a place defined by a polygon with its coordinates, a toponym, or a Boolean combination of toponyms. This jurisdiction must be included within the jurisdiction of the rule set.

Then a list of relevant geographical objects and possibly Boolean conditions will follow.

The part "implication" should indicate whether this rule is imperative, fuzzy, frequent or good type practice.

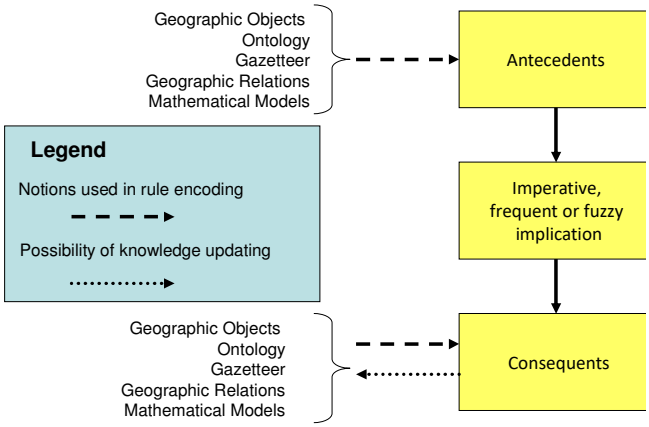


Figure 8. Geographic knowledge involved in the geographic rules.

As "Consequent", there may be the change of geographic objects geometrically or semantically and launching of actions.

In addition, for a rule, it would be interesting to add metadata and an explanatory text. The role of metadata will primarily be to mention the origin of the rule, legal type, data mining, good practice, etc. as well as the date and the name of the editor. As the explanatory text, this will be the text that will be printed when the user may wish to reconstruct the path of reasoning (traceability of the result).

VI. Concluding Remarks

The purpose of this paper was to illuminate the notion of geographic rules by examining examples in order to capture their semantics. Unlike the rules of management in enterprises, we have tried to show the importance of space and the difficulties it could lead. Ultimately, the strong elements of the semantics of this type of rules were extracted, which allowed us to develop a first model. But we need other examples showing for instance other aspects which were not discovered yet. Then, we need to build an inference engine capable of integrating and reasoning with this type of semantics.

Therefore, it is now possible to give a definition of a geographic rule that can be set as an imperative or modulated implication (frequent, desirable, etc.) involving either places or geographic objects, or both.

Another track will be to take account of temporal aspects to describe the rules of evolution of geographic objects such as shape changes (forests and deforestation, urban sprawl, dissemination, flood, etc.). Also 3D should be included.

But first and foremost, it is necessary to continue this analysis in order to introduce other cases and thus enrich the semantics of the geographic rules. Then, it will be possible to propose a robust, consistent and effective formalism for representing the geographic rules and enable them. It will also need to define the precise specifications of the actions to be undertaken for the treatment of the modulated implications.

References

[1] Batty. M, Yeh T. (1991) The promise of expert systems for urban planning, Computers, Environment and Urban Systems, Volume 15, Issue 3, 1991, pp. 101–108.

[2] Biasotti S., Cerri A., Patané G., Spagnuolo M. (2015) Feature Extraction and Classification. In Heterogeneous Spatial Data: Fusion, Modeling, and Analysis for GIS Applications, Edited by Patané G., Spagnuolo M., Morgan & Claypool Publishers, California, pp. 57-81. Under press.

[3] Boley, H., Paschke, A., & Shafiq, O. (2010) RuleML 1.0: The Overarching Specification of Web Rules. In Semantic Web Rules: International Symposium, RuleML 2010, Washington, DC, USA, October 21-23, 2010, Proceedings (Vol. 6403, p. 162). Springer Science & Business Media.

[4] Borges, K. A. V., Davis Jr., C. A., Laender, A. H. F. (2002) Integrity Constraints in Spatial Databases. In: Doorn, J. H., Rivero, L. C. (Org.) Database Integrity: Challenges and Solutions. Hershey (PA), USA: Idea Group Publishing, 2002, pp. 144-171.

[5] Bravo L., Rodríguez A. (2012) Formalization and Reasoning about Spatial Semantic Integrity Constraints. Data & Knowledge Engineering 72 pp. 63- 82.

[6] Dietz J.L.G., (2008) On the Nature of Business Rules. In Advances in Enterprise Engineering, Springer Verlag Lecture Notes in Business Information Processing (10) pp 1-15

[7] Graham I. (2006) Business Rules Management and Service Oriented Architecture: A Pattern Language. London, John Wiley.

[8] Jain K, & Payal (2011) A Review Study on Urban Planning and Artificial Intelligence, International Journal of Soft Computing and Engineering (IJSCE), ISSN: 2231-2307, Volume-1, Issue-5, November 2011

[9] Kim, T.J., Wiggins, Lyna L., Wright, J.R. (Eds.) Expert Systems: Applications to Urban Planning, Springer-Verlag, New York (1989)

[10] Laurini, R. (2014) A Conceptual Framework for Geographic Knowledge Engineering, Journal of Visual Languages and Computing (2014), Volume 25, pp. 2-19.

[11] Laurini, R. (2017) Geographic Knowledge Infrastructure for Territorial Intelligence and Smart Cities. ISTE-Wiley. 250 p. Forthcoming Spring 2017.

[12] Morgan T. (2008) Business Rules and Information Systems: Aligning IT with Business Goals. Addison-Wesley.

[13] Ross R. G. (2011) More on the If-Then Format for Expressing Business Rules: Questions and Answers, Business Rules Journal, Vol. 12, No. 4 (Apr. 2011), URL: http://www.BRCommun 2002ity.com/a2011/b588.html.

[14] Salleb-Aouissi A., Vrain C., Cassard D. (2015). Learning Characteristic Rules in Geographic Information Systems. N. Bassiliades, G. Gottlob, F. Sadri, A. Paschke, D. Roman (Eds.). Rule Technologies: Foundations, Tools, and Applications, 9th International Symposium, (RuleML 2015), Aug 2015, Berlin, Germany. Springer, 9202, 2015, Lecture Notes in Computer Science.

[15] Servigne S., Ubeda T., Puricelli A., Laurini R. (2000) A Methodology for Spatial Consistency Improvement of Geographic Databases, GeoInformatica 4:1, pp. 7-34.

[16] Varadharajulu P., West G., McMeekin D. Moncrieff S., Arnold L. (2016), Automating Government Spatial Transactions. International Conference GISTAM Conference, Roma, April 2016.

Automatic Generation of an Interactive, Real time, Web-based Maps from Sensor-based Geographic Databases

Zina Bouattou, Hafida Belbachir
Computer Science Department
Oran University of Science and Technology
Oran, Algeria
zina.bouattou@univ-usto.dz, h_belbach@yahoo.fr

Robert Laurini
LIRIS, INSA de Lyon, University of Lyon, France
and Knowledge Systems Institute,
Illinois, USA
Robert.Laurini@insa-lyon.fr

Abstract — In recent years, we have seen a growing number of location-based sensors used to measure increasingly diverse phenomena. Therefore, we consider the increasing need to efficiently manage the huge amount of data collected from located wireless sensor networks (WSN). In some phenomena, it is fundamental to generate maps in real time. Data can be grouped into web-cartography for diverse purposes; so finally, we develop a web-based mapping on demand. In this paper, in addition to cartographic functionalities mentioned above, we are facing another challenge. Indeed, we encounter the problem of avoiding on-the-fly overlays of multiple visual layers in order to get a result that is easy to interpret due to visual interference between graphic elements. The present paper describes an approach to the cartographic visualization of geo-data regularly acquired by WSN and exhibits a prototype portal based on an open source platform for integrated visualization as visual realtime summaries. An application in meteorology is proposed.

Keywords-component; real-time database; sensors, real-time cartography; interactive visualization; geovisualization, semiology.

I. INTRODUCTION

In the last decade, a new concept has emerged known as "Digital Earth" for which one is seeking ways to ensure access to information resources and to deliver geo-referenced knowledge to everyone. To reach this goal, many problems need to be addressed. One of the key issues is the development of novel methods coupling big data arriving in real time from sensor networks and the appropriate visualization in a way that allows users of its rapid and correct understanding. As visual presentation and analysis of these data are currently a very promising research topic; the scope will be not to determine the mapping once at all, but more especially to calculate it in real time in order to visualize evolution and to understand information in space and time by means of visual summaries.

In addition, much of the research effort in geovisualization research has made to improve dynamic, interactive maps, especially for broadcast on the Internet [1]. The number of applications has grown exponentially with the advent of application programming interfaces (APIs), such as Google Maps API[1], OpenLayers API[2], OpenStreetMap[3]. Further research actions in exploratory geovisualization have resulted in significant efforts to understand the use and improve the usability of interactive applications for the Web-based maps [2, 3]. However, many current systems only allow static layout for relatively large data sets display. In the current web cartography, tools will require unique solutions for navigation, querying, viewing and interpretation of millions data in space and time.

Moreover when it comes to real time geographic data coming from the sensor network, here the complexity of geovisualization increases further. In addition, existing cartographic design principles must be adapted to cope with the production of on-the-fly maps associated with large volumes of data.

In this paper, our challenge is multiple:

- Because of the problems caused by data entry and processing time data streams coming from sensors networks, we are motivated to introduce a dynamic system that will collect and input data automatically into a GIS system, to process and to visualize these data on a real time web-cartography.
- Graphic semiology for real-time environments: The integration of real-time data in interactive cartographic visualization has gained little attention in research to date, many maps (especially digital versions) are of poor cartographic quality containing huge amounts of data (e.g. synoptic maps), despite existing guidelines. Accordingly, these maps are difficult to understand and interpret what hinders the flow of information and communication of important phenomena available, so, how can we improve the presentation and cartographic visualization communication?
- In some kind of maps, usually called atlases, only one variable is mapped separately and often two of them are presented together. But sometimes there is an attempt to

[1] https://developers.google.com/maps/
[2] http://openlayers.org/
[3] https://www.openstreetmap.org/

DOI reference number: 10.18293/DMS2016-022

show three or four variables with a result that is difficult to interpret due to the interference between visual graphics, so the question is: how in a single map computed in real-time, to visualize multiple variables in a very readable and understandable way?

In this paper, we will deal with these questions and introduce a new approach method of visualization of sensor data. The rest of the paper is organized as follows. Section 2 presents related works. Section 3 presents the methodology of our system. Section 4 we present our preliminary result and discussion. We conclude our paper in section 5 and suggest some further perspectives. Along the paper, an example is used, based on meteorological data.

II. STATE OF THE ART

In this section, we briefly discuss related works required to understand our research.

A. Sensor network

Sensor network consists of an autonomous node set, battery-powered, randomly distributed in a geographic area that defines the territory of interest for the captured phenomena; each of these nodes has the capability to collect and route data either to other sensors or to some gateways [4] (see Figure 1).

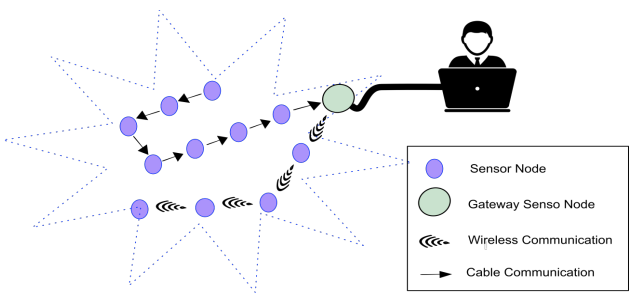


Figure 1. Sensor networks architecture.

Data streams coming from sensor networks are defined as "sequences of ordered data (usually by arrival time), continuous and real-time" [5]. They are different from conventional data, characterized by the spatial dimension of data which determine the sensor position, and the time dimension of the data which determines the moment of shared sensors measurement. Both are carriers of information and play a crucial role in the information synthesis. However, the real-time data streams are not easy to retrieve and to process. Few software products and applications dealing with such data due to their continuous nature and volume and all in real-time, manage such data in a workable set of constraints and challenges.

B. Cartography, GIS, geovisualization

Cartography is an ancient discipline developed from the practice of making maps, combining science, and complex techniques in an effort to portray the world accurately and effectively convey information to the map-reader [1]. As a GIS (acronym of Geographic Information Science), was developed much more recently as software systems designed to perform a wide range of operations on geographic information. GIS is supported by progress from various disciplines such as geography, surveying, engineering, space science, computer science, cartography, statistics, and so on. Cartography is both an art and a science. Many researchers found the traditional term cartography too limited as a description of this new, much richer world, and begin to describe their field as geovisualization. Geovisualization (GVis) [1] is considered as a means of representing spatial information visually in a way that allows people to explore, synthesize, refine, analyses, and communicate conclusions and ideas. Geovisualization is much more than cartography since it includes other methods to represent the geographic reality for decision-makers.

C. Related Works in real-time Web-cartography

Integrating web cartography has been studied since 1997 [6]. Several techniques have been added to maps, making them more interactive, dynamic, multimedia and improving the access to information (accessibility, timeliness, authenticity). These techniques fully exploiting its potential must be addressed for the need in terms of more complete communication and effective and improved cartographic analysis. The design of the web cartography maps is constituted by a set of sophisticated visualization functions which allow users to interact with maps through a GUI (Graphical User Interface) for example by increasing the level of details by zooming, remove less important elements by filtering, GUI must be laid out clearly and functions must be easily identifiable and manageable to ensure an uncontained use and intuitive system [7].

Another starting point is in the information visualization topic, several authors [8, 9] adopted the Visual Information-Seeking mantra stated by Ben Shneiderman [10], namely "Overview, Zoom and Filter, Details on Demand" as a starting point to create a good visualization; it is a well-known visualization paradigm which encompasses several visual design guidelines and provides a general framework for designing information visualization applications:

- Overview: gain an overview of the entire collection;
- Zoom: zoom in on items of interest;
- Filter: filter out uninteresting items;
- Details-on-demand: select an item or group and get details when needed.

Interactive functionality and web-based technology for an expert tool in the field of natural phenomenon management has been identified in recent research, in [3] authors have developed a web-based application for real-time visualization of hydrological data. This application provides tools to control interactively, trace, and compare available information. Another research project focuses on the methodology and scientific issues involved in real-time cartographic in operational hydrology field [11]. This project aims at automating mapping steps applied to real-time data measurement from sensors. To serve as a proof of concept and demonstration purposes, an information system to map in real time hydrological data was established. It consists of a space

base in real time and a user interface based on Web mapping. This kind of application allows real time control of all operations and also allows synchronization with time and especially in the field of early warning, for a discussion of the strategy to introduce real time in the cartographer the reader can refer to [12].

D. Semilogy and cartography

Cartography is the result of a structured model, organized; the final quality resulted from the inclusion of simple rules of graphic semiology and map design. They follow the graphic semiology developed by Bertin [14]. Web cartography is associated with technical restrictions and specific semiotic requirements. Signs, graphic variables, color, context, density are all aspects to be considered in the cartography.

Since we only assessed 2D maps, Jacques Bertin's definition of the characteristics of point [13], line and area symbolization served as a basis. These six visual variables: color hue, color value, shape, size, orientation, and texture have been actively discussed and extended by several authors. However, we only included variables that were present in at least one of the assessed cartography. Bertin's six visual variables were consequently extended with color saturation added by MacEachren [14] and transparency added by Wilkinson [15]. These variables were analyzed in order to assess which variables are used for natural hazard visualizations and what values they are assigned. Figure 2 gives an overview of the visual variables.

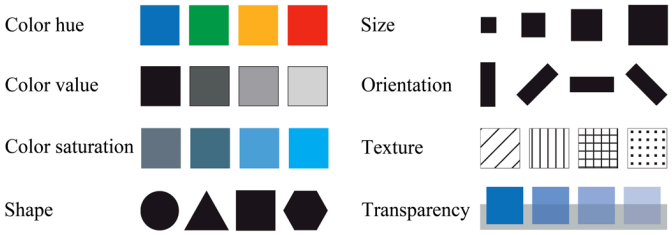


Figure 2. Overview on the visual variables [21]

Pang [16] addresses the issue of the uncertainty associated with data on natural hazards in detail and presents potential methods to visualize the uncertainty of natural hazards such as the application of blurriness, transparency or shake, using the tint color, saturation, or value, a grid overlay which is modified according to the uncertainty values, the drawing of the contour lines, the variation of the thickness, brightness or connectivity symbolization, the use of glyphs, histograms or plots.

Color is a visual variable widely used in large-scale cartography. Different color shades allow different thematic distinguished. Generally colors ranging from blue are often used for water depths of illustration, but also to show a decrease amount. The red colors, are often used to show a risk, a hazard, but also an increase of the amount. While green is often used for forestry, tan for deserts. Color choice should be made by carefully because it is an important variable to create visual attention [19]. A bad choice of color can confuse user map [20].

More than 90% of cartographers used color hue for the symbolization area in maps, the colors are preferred in the realization of a large-scale map [21]. Although the texture is a visual variable that would be well suited to solve the map overlay multivariate problems, only some of the maps contain the variable analyzed visual texture. If more than one layer of thematic data is displayed, the textures can superimpose and lead to the omission of important information thanks to the effect of laciness [18]. In addition, if the texture is carefully chosen and applied, it may seem chaotic and disturbing the overall balance of the map.

III. METHODOLOGY : AN GIS BASED WSN DATA VISUALIZATION SYSTEM

In this section, we discuss how challenges described above are solved in the proposed methodology. Specifically we first briefly describe the details about the manner we build GIS visualization systems with WSN geo-data. We then present the perceptual basis for our design map and describe its implementation in a public weather display in comprehensive adopted the Visual Information-Seeking mantra stated by Ben Shneiderman [10]. Then because our prototype WSN is deployed to monitoring we chose to superpose multiple layers of visualisation to inform the developments made by Plaisant, Wilkinson, and Bertin [14, 15, 16]. We finish with a discussion of the implications of those results.

A. Used data and test area

Real-time data set processed in the project comes from OpenWeatherMap API[4]. This API provides access to weather sensor stations that are directly connected to the Internet and enables us to collect high frequency updates from potentially hundreds of thousand sensors deployed over a larger area. We used altogether 160 stations distributed in the region of France. With such a data format, we can potentially use a variety of meteorological data to describe different weather phenomena, their distribution organization in space. The temporal resolution is chosen as fifteen minutes. Active sensors acquire data stream for each climate, density variable, perhaps a few values such as: temperature, humidity, pressure, speed and direction of wind, the presence of clouds and precipitation.

B. Real-time workflow and functions

Figure 3 gives an overview workflow which shows the various components used in the system and how they are connected to each other. To view the sensor data collected from our WSN deployed, we need three steps:

- Geo-data collection,
- Geo-data preprocessing and storage,

[4] http://openweathermap.org

- Geo-data display in web-cartography (semiology...)

1) Geo-data collection*: URL to database*

Each of the stations contributes from the *OpenWeatherMap API* to an *XML* file on current meteorological conditions; this file is updated every time the station sends a new set of observations. There is currently no direct method to retrieve these conditions concerning several weather stations; they are individually accessible via an *URL* that contains the *ID* of the station of interest. To create a complete set of meteorological observations for France, which can then be analyzed, a small Java program was developed to collect data streams from several stations iteratively each fifteen minutes, and the observations are then stored in a PostgreSQL/PostGIS [5] database.

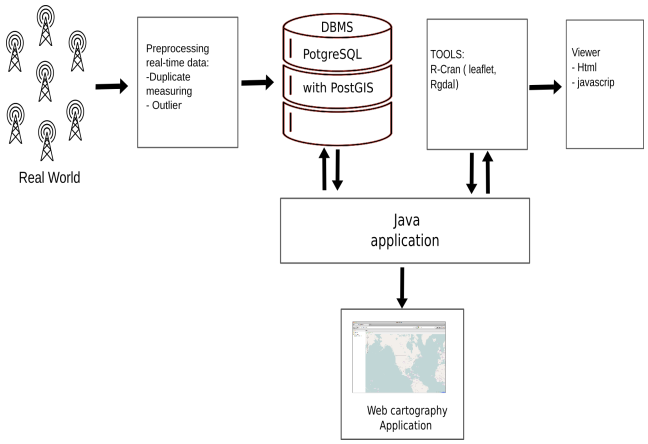


Figure 3. System architecture and used technologies for the workflow.

PostGIS extension can provide support to the database for storing multiple geographic coordinates. The reason we chose PostgreSQL is because it is open source, popular and supported by a wide variety of systems (See Figure 4). A JDOM[6] API is dealing the browse of the returned XML[7] document and retrieves the data to immediately decipher the response returned by the service (OpenWeatherMap API).

2) Geo-data preprocessing and storage

Before the insertion into the database will occur, data are checked and cleaned (e.g., duplicate measuring time, type of data such as text or number, outliers). If the data fail to meet these conditions, either no data or pre-set values are stored. So we had to consider a degree of confidence for data in real time when used for filtering and comparison with the historical data.

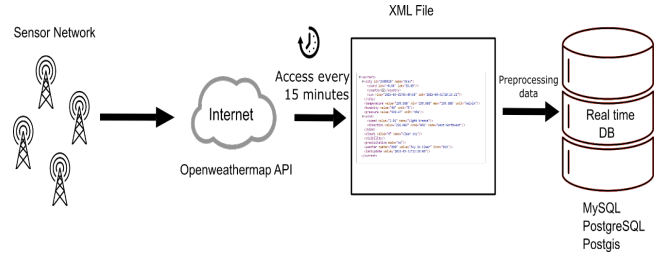
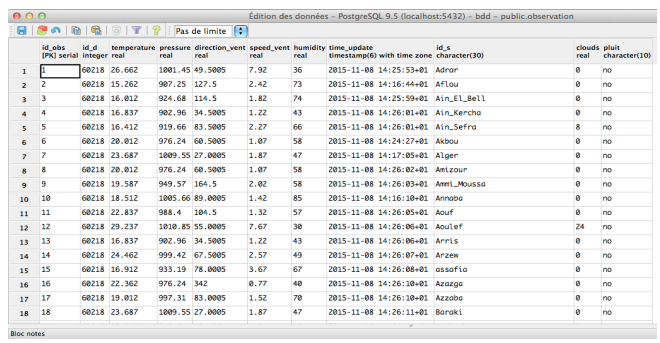


Figure 4. Real time acquisition geo-data process

We use R-project to create a web-cartography and geographic information using Leaflet[8] and Rgdal[9] package. The first reason or choosing R is that its open source software and freely available. And the second was that possible to generate directly a web-map from R.

3) Geo-data display in web-cartography

After the work steps between the extraction and the backup of the database, the data is automatically converted into graphic formats so that they can be presented in the cartographic interface based on the Web. In this section, we present the details of how we present the geographical data. These include how to display the map, the different layer data in web-cartography application, semiology, the visual value.



	id_obs [PK] serial	id_d integer	temperature real	pressure real	direction_vent real	speed_vent real	humidity real	time_update timestamp(6) with time zone	id_s character(30)	clouds real	pluit character(10)
1	1	60218	26.662	1001.45	49.5005	7.92	36	2015-11-08 14:25:53+01	Adrar	0	no
2	2	60218	15.262	907.25	127.5	2.42	73	2015-11-08 14:16:44+01	Aflou	0	no
3	3	60218	16.012	924.68	114.5	1.82	74	2015-11-08 14:25:59+01	Ain_El_Bell	0	no
4	4	60218	16.837	902.96	34.5005	1.22	43	2015-11-08 14:26:01+01	Ain_Kercha	0	no
5	5	60218	16.412	919.66	83.5005	2.27	66	2015-11-08 14:26:01+01	Ain_Sefra	8	no
6	6	60218	20.012	976.24	60.5005	1.07	58	2015-11-08 14:24:27+01	Akbou	0	no
7	7	60218	23.687	1009.55	27.0005	1.87	47	2015-11-08 14:17:05+01	Alger	0	no
8	8	60218	20.012	976.24	60.5005	1.07	58	2015-11-08 14:26:02+01	Amizour	0	no
9	9	60218	19.587	949.57	164.5	2.02	58	2015-11-08 14:26:03+01	Ammi_Moussa	0	no
10	10	60218	18.512	1005.66	89.0005	1.42	85	2015-11-08 14:16:10+01	Annaba	0	no
11	11	60218	22.837	988.4	104.5	1.32	57	2015-11-08 14:26:05+01	Aouf	0	no
12	12	60218	29.237	1010.85	55.0005	7.67	30	2015-11-08 14:26:06+01	Aoulef	24	no
13	13	60218	16.837	902.96	34.5005	1.22	43	2015-11-08 14:26:06+01	Arris	0	no
14	14	60218	24.462	999.42	67.5005	2.57	49	2015-11-08 14:26:07+01	Arzew	0	no
15	15	60218	16.912	933.19	78.0005	3.67	67	2015-11-08 14:26:08+01	assafia	0	no
16	16	60218	22.362	976.24	342	0.77	40	2015-11-08 14:26:10+01	Azazga	0	no
17	17	60218	19.012	997.31	83.0005	1.52	70	2015-11-08 14:26:10+01	Azzaba	0	no
18	18	60218	23.687	1009.55	27.0005	1.87	47	2015-11-08 14:26:11+01	Baraki	0	no

Figure 5. Overview of the archive data collection

In this paper, we have used four different layers to complete the map application. There are several basic components that are implicated in the development of web-cartography: basemap; thematic layers; additional information; interactivity; legend.

a) Basemap

There are eight types of predefined maps currently supported by ESRI. They are streets, satellite, hybrid, topography, grey, oceans, national-geographic, and OpenStreetMap (OSM).

[5] http://www.postgresql.org/
[6] http://www.jdom.org/
[7] Extensible Markup Language

[8] JavaScript library for interactive maps
[9] Geospatial Data Abstraction Library

```
library(leaflet)
map <- leaflet()
map <- addTiles(leaflet())
map <- fitBounds(map = map,
                 lng1 = -4.445833,
                 lat1 = 40.50306,
                 lng2 = 9.484722,
                 lat2 =  50.56417)
# Map Size
map$width <- 1000
map$height <-700
map<- addDrawToolbar(m)
map
```

Figure 6. Extract R-code basmap example

In our application we have used OSM as a basemap. OSM is used because it is an open and freely available database of geographic data. Moreover, it provides more desirable locations and easier traceable views than other available maps. An example in R is given in Figure 6.

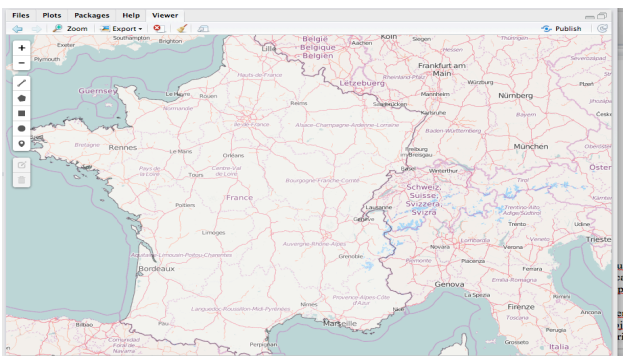


Figure 7. France Basemap with OpenStreetMap(OSM).

Here is the map built a basemap map inside the given HTML container. The scope of the map is the size of the container. The manufacturer Map includes a basemap name as the first parameter. Therefore, we use the fitBounds center setting, this ensures that, when a user opens the application, it is automatically centered on a desired location, so that the user does not need to know where is located, we define the map size by size map settings, default map has a contextual option that displays a minimum value related to the clicked point. As the default context menu cannot serve our purpose, we used to popup to display additional information.

b) Thematic layers

Visualization is important when working with data from sensors. According to Plaisant [17] the biggest problem when designing visualizations is mounting and the combination of visualization to the user's wishes. Try to view the data; different views of the same data set should be available. Recall our second objective; in real time, it is important particularly in the meteorological field when decisions must be taken in short terms, within hours or minutes. Such situations leave no time for the reader, it loses a lot of time between several maps to see several phenomena, but if in a single maps several phenomena without presented bunk and interactively or it can compare and interact with the map that will be more interesting to the reader for taking immediate decision. According to [21], it is impossible for humans to examine the information and understand the content and then discover and combine it in a short time under the pressure of time in disaster management.

Overlaying many of the different thematic layers is still a challenge in cartography: it is the central component of the application. To construct these layers, the database connection operations have been used to make the database available to R. When properly selected, it makes it more comfortable for a user to work with data and data can be understood more quickly and easily. The weather maps generally show the information using the contour intervals resulting from an interpolation method. Our mapping solution is to use visuals three channels; a color channel, a texture channel, and a symbol channel, to separate the perceptual variables and make them independently readable. Remember that the second objective of this paper is to provide visualization more easily and a method of combination that clarifies understanding, but not to obscure it. To this end, three different visualizations are used.

We chose to map the atmospheric pressure interval of color codes because it is so pervasive in modern weather visualizations whereas the wind speed is usually represented by proportional symbols. We added the digital text to provide additional wind information for the wind direction which is given by symbol orientation. Finally, we have chosen to represent temperature with texture with the use of contours, because it was the only channel and we found later that it provided a great opportunity to mix the use of texture with the use of contours, which are commonly used to represent the pressure in the weather maps and make abrupt changes in pressure visually prominent.

c) Generalization

Modern IT environments allow assessments of high spatial resolution cartography, which results in huge amounts of data. These raw data must be pre-processed before it can be viewed to avoid chaotic and overloaded maps. For these generalization techniques to be applied, it includes interpretation, selection and omission of facts available. To keep the balanced mapping of available information can be neglected or given a lower level of importance. An example is the smoothing of raster data so that the edges are rounded and can be integrated in the maps without being visually striking (see Figure 8).

Another important task is filtering the raw data before they are presented to the map-readers. The greatest responsibility takes over this task; an easier map can be read and interpreted by users

a) Interactivity

Zoom, filter, panoramic and layer visibility management are all functionalities that provide interactivity with

cartography. A more sophisticated mapping portal offers additional features such as tooltips or windows with information on attribute values, measure distances, search functions, uncertainty values and printing options.

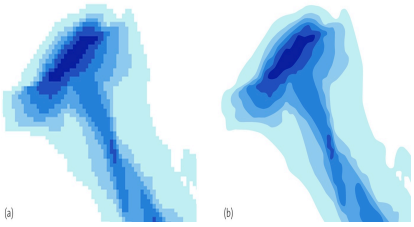


Figure 8. Raw raster data (a) can be visually striking and should be smoothed in order to obtain a natural appearance (b) [21].

b) Legend

The legend is a key that contains information on the thematic areas present; it supports the orientation of the map-reader. If the legend is not clearly visible or if ignored, by the map-readers, they cannot decrypt the contents of the map can lead to misinterpretation.

IV. PRELIMINARY RESULTS AND DISCUSSION

The outcome of implementing these methods is illustrated through a graphical user interface (Figure 9), it displays a simple form of the map contains meteorological data in real time that are automatically modified, processed, and displayed in an interactive web cartography and meets the basic principles of readability and comprehension, map illustrates how geographic data can be viewed by vector maps where various geometric and pictorial map symbols are presented. The expected temperatures are displayed as an interpolated surface filled with a visual value textured where the variation of the thickness proportional to the temperature value associated with the caption displayed in the auxiliary window (second legend right).

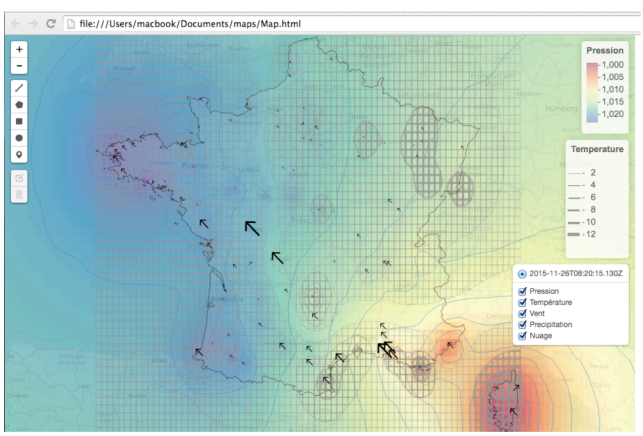


Figure 9. Web cartography created in real time

During writing of this paper, since we were unable to find previous examples using visualization isolines filled proportional texture, it was an opportunity to integrate and test in our prototype. As pressure is represented by isolines filled by the color level as depicted in legend right up. The wind speed is usually represented by proportional symbols; we added the digital text to provide additional wind information for the wind direction is given by the symbol orientation (an example is shown in Figure 10).

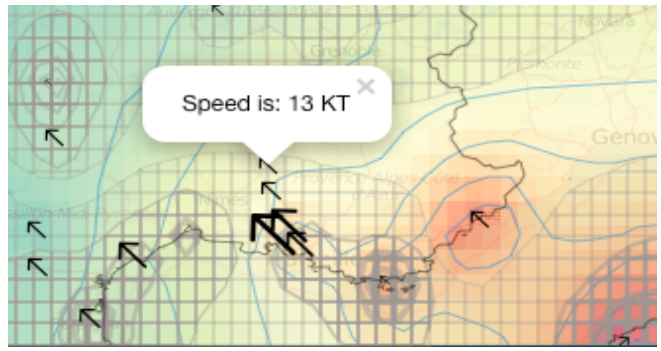


Figure 10. Popup window showing wind speed information

In our current implementation, the user has the ability to customize data. However, the user can interact with the map and change the view by interacting with the interactive menu (right down), followed by the date of the last updates as shown in Figure 10. In addition, spatial navigation controls such as zoom in/out, panoramic visual and re-centering are available, so users must be able to zoom in/out, and even a map of the edition menu (left of the interface).

Moreover, when users click on any location on the map, he opens a window on the map corresponding to the particular sensor or estimated values of adjacent sensor for more information. This is illustrated in Figure 11. At present, the service is not fully complete; however, it follows all the major concepts and supports mapping data arrives in real time from the sensor distributed over a territory.

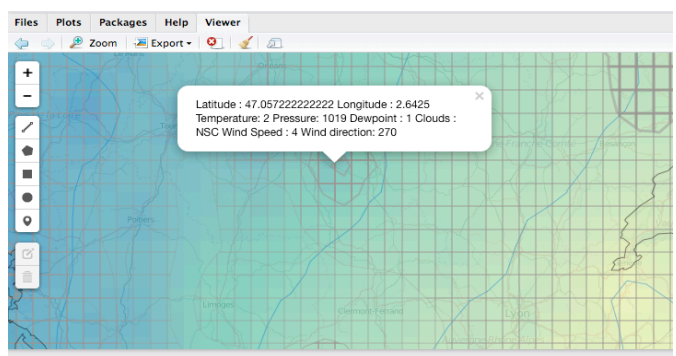


Figure 11. Popup window showing additional information

For exploration and visualization of measured sensor data, an interactive web map was developed; application complies with the basic design principles outlined above. Namely the multiple views and tasks and principle were followed as (overview, zoom and filter, and details on request) principles were followed. In this paper, our objective was twofold; firstly, we intend to have deeper processes parts of the

automated process of digital cartographic modeling and its implications for web-based interactive cartography.

On the other hand, we expect progress in the area of the temporal and real time GIS mapping, which is today a challenge that concerned many researchers. With our prototype, we would like to provide an application that can significantly reduce the work offline by automating much of the cartographic design process.

IV. CONCLUSION AND OUTLOOK

Sensor data are typically in the form of digital values; therefore, the understanding of processes or analysis is not trivial. Visualizations can be an interesting solution which is a complex process; take the digital data, analyze, use methods of extraction and aggregation, and then give the user a graphical interface represents a challenge. The ultimate goal is to provide automated way, real-time visualizations from the raw data of distributed measurement while respecting rules and modern cartographic standards.

In this paper, we present our first initial research efforts to store and manage and visualize big sensor data in a WSNs prototype based on GIS. By focusing on the cartographic visualization and friendly interaction to facilitate understanding the information, we present a web cartography application that provides responses to data requests effectively. We prove by our workflow, as the sensor data processing to achieve the visualization is possible on the fly. This allows users to visually relate the phenomenon represented by superimposing layer, it is the second objective of the research presented here; develop and implement strategies to symbolize appropriately aggregated data to support the visual interference. Our works are still in progress and research can be extended to include the following future work:

- First, we intend to add more functionalities as time navigation; we also plan to add alerts for certain types of data that automatically notify an interested user group.
- In order to establish rules of best practice in this area by facing this objective, a psycho-cognitive test can be presented as a questionnaire to validate the choice of the visual vocabulary order to interpret the results of the application of a simple and correct way.
- The visualization presented in real time requires a fully automated processing and no error due to non-possibility of verification data map before displaying. The main difficulty coming from the complexity of the workflow is the quality of data that is unknown. This problem was solved by a simple checking of the validity of real-time values on-the-fly before viewing, but we think that we still need more thorough research in our future work to solve this problem.

REFERENCES

[1] A.M. MacEachren and M. Kraak, "Research challenges in geovisualization", Cartography and Geographic Information Science, vol 22(1), 2001, pp. 3-12.

[2] T. Bhowmick, A.L. Griffin, A.M. MacEachren, B.C. Kluhsman and E.J. Lengerich, "Informing geospatial toolset design: Understanding the process of cancer data exploration and analysis". Health and Place, 2008.

[3] C. Lienert, C. Edouard R.Weingartner and H.Lorenz, "Real-Time Cartography in Operational Hydrology", Cartography & Geographic Information Systems, vol 36, 2009, pp.45-58.

[4] C. David, E. Deborah and S. Mani, "Guest editors' introduction: Overview of sensor networks", Computer, vol 8, 2004, pp: 41-49.

[5] B. Babcock, S. Babu, M. Datar, R. Motwani, J. Widom, "Models and Issues in Data Stream Systems", Proc of the Twenty-first ACM SIGMO D-SIGACT-SIGART Symposium on Principles of Database Systems, Madison, WI, USA. (1-58113-507-6), 2002, pp.1-16.

[6] M.P. Peterson (Ed.), "Maps and the Internet", Amsterdam: Elsevier. 2003

[7] R. Sieber, C. Schmid and S. Wiesmann, " Smart Legend – Smart Atlas!", Proceedings of the 22nd International Cartographic Conference ICC, A Coruña, Spain. 2005

[8] D. De Chiara, V. Del Fatto, R. Laurini, M. Sebillo, G. Vitiello, "A Chorem-Based Approach for Visually Analyzing Spatial Data", Academic Press: In Journal of Visual Languages Computing, vol 22(3), 2011, pp.173-193.

[9] N. Andrienko, G. Andrienko, "Exploratory Analysis of Spatial and Temporal Data", Berlin, Germany: Springer, Springer – Verlag, Berlin Heidelberg, pp. 703 . 2006.

[10] B. Shneiderman, " The eyes have it: a task by data type taxonomy for informationvisualizations", In Visual Languages, 1996. Proceedings., IEEE Symposium, 1996, pp: 336-343.

[11] C. Lienert, R. Weingartner and L. Hurni, "An interactive, web-based, real time hydrological map information system", Hydrological Sciences Journal, vol 56(1), 2011, pp.1-16.

[12] H. Moellering, "Strategies of Real-Time Cartography". The Cartographic Journal, vol 17(1), 1980, pp. 12-15.

[13] J. Bertin,"Sémiologie graphique les diagrammes – les réseaux - les cartes". Paris, France: Mouton. 1967.

[14] A.M. MacEachren, M.J. Kraak, "Exploratory cartographic visualization: advancing the agenda", Computers & Geosciences, vol 23(4), 1997, 335-343.

[15] L. Wilkinson, "The grammar of graphics". Satistics and computing, Springer, New York, 1999.

[16] A. Pang, "Visualizing uncertainty in natural hazards". In: Bostrom, A., S.P. French, and S.J. Gottlieb (eds). Risk assessment, modeling and decision support. Springer, Berlin Heidelberg, 2008, pp.261-294.

[17] C. Plaisant, "The challenge of information visualization evaluation". In Proceedings of the working conference on Advanced visual interfaces, ACM, 2004, pp. 109-116.

[18] T. Watababe and P. Cavanaugh, "Texture laciness: the texture equivalent of transparency", Perception, vol 25, 1966, pp.293–303.

[19] J.M. Wolfe and T.S. Horowitz, "What attributes guide the deployment of visual attention and how do they do it?" Nature Reviews Neuroscience, vol 5(6), 2004, pp.495-501.

[20] P.R. Keller and M.M. Keller, "Visual cues: Practical data visualization, IEEE Computer Society Press. 1994.

[21] M. Kunz and L. Hurni, "How to enhance cartographic visualisations of natural hazards assessment results". The Cartographic Journal, 48(1), 2011, pp. 60-71.

VeCVL: A Visual Language for Version Control

Nathan W. Eloe, Denise M. Case
School of Computer Science and
Information Systems
Northwest Missouri State University
Maryville, MO 64468, USA
{nathane,dcase}@nwmissouri.edu

Jennifer L. Leopold
Department of Computer Science
Missouri University of Science and Technology
Rolla, MO 64468, USA
leopoldj@mst.edu

Abstract

Version control systems (VCS), such as Subversion and Git, are pervasive in industry; they are invaluable tools for collaborative development that allow software engineers to track changes, monitor issues, merge work from multiple people, and manage releases. These tools are most effective when they are a part of a developer's habitual workflow. Unfortunately, the use of these powerful tools is often taught much later in a developer's educational career than other tools like programming languages or databases. Even an experienced student's first experience with version control can be unpleasant. In this paper, the authors analyze the workflow of two common Version Control Systems with different version controls (Subversion and Git) to build a common visual language for these systems (Version Control Visual Language, or VeCVL), and show that the same visual language applies to other version control systems.

Keywords- *computer science education, education technology, pedagogy, version control, visual language*

1 Introduction

Mastery of version control is vital to a developer in modern software engineering; the scale of current projects requires collaboration from teams of developers to correctly and efficiently implement, maintain, and release software that solves real-world problems. Some projects have immense code bases, some require experts in different areas, and some (such as the Linux kernel) fall into both categories.As with any tool in a software engineer's repertoire, the VCS is most effective when used frequently; however it is one of few tools that can be detrimental to the success of a project unless its use is habitual. Poor or incorrect use of a debugger can slow down development, but a developer is unlikely to destroy a project by "jumping into" a function definition instead of "jumping over" it. Incorrectly using a source control tool can destroy other developers' changes, modify history, or make it difficult to tag and release a new version of the software. To realize the full potential of version control, the system must be used frequently, with small changes.

Unfortunately, these tools are often taught late in a student's educational career. They may be introduced in a software engineering course (often not one of the first courses a student takes), and may not be reinforced in future classes. These systems can also be difficult to use; for all their power and flexibility, they can be unfriendly and require students to make drastic changes to their normal development workflow.

This paper, in conjunction with concurrent research [1], is an investigation into embedding the use of version control within pedagogy. To simplify interaction with these tools, a visual language that supports the basic operations of common version control systems is presented. It aims to reduce the severity of the learning curve these tools have and ease the transition between different systems. Additionally, an analysis of the initial implementation of VeCVL is presented.

2 Background and Related Work

In this paper, we build on research in the areas of instructional technology and pedagogy, visual languages, and version control systems.

2.1 Instructional Technology

Software engineering and development can be challenging, and a variety of pedagogical approaches have been developed to assist with the learning process [4]. *Scaffolding* is a popular instructional technology that supports beginning students in accomplishing challenging tasks [11]. As the student becomes more proficient, the scaffolding can be

DOI reference number: 10.18293/DMS2016-027

gradually removed (through a process called *fading*) until the support is no longer needed. In this work, we propose to provide a visual language for new software engineers as a scaffolding mechanism to facilitate learning and the transition to native VCS environments.

2.2 Visual Languages

To facilitate the introduction and application of version control tools in developer's workflow, we build on the work done in the area of visual languages. A visual language is a *system of communication using visual elements rather than letter strings* [10]. The introduction of useful visual indicators has been shown to improve cognition by leveraging our human ability to see patterns and trends [5]. Our work develops icons for operations in the work flow and follows conventions for visualizing relating or overlapping features in a common way as recommended by Jones [6].

2.3 Version Control Systems

Readers interested in the history of modern version control systems (VCS) should should refer to [9]. Of interest in this paper is the development of two different paradigms of VCS: Centralized and Distributed.

2.3.1 Centralized VCS

A centralized VCS works on the principle of having a single shared repository that accepts code; multiple developers pull the latest working version from the shared repository, make changes, and then synchronize their changes to the repository [2]. This workflow is simple to understand and widely used.

One consequence of this workflow is race conditions: if two developers are working on the same code base, and the first developer pushes changes to the central repository, the second developer must first update their local copy with the changes from the global repository before they can upload their changes.

One widely adopted Centralized VCS (CVCS) is Apache Subversion, or SVN. SVN is an open source, mature CVCS that is a fully semantic successor to CVS, and is used to host code for many open source projects and groups, including WebKit and the Apache Software Foundation [2]. The basic work cycle in SVN is outlined in [8] and visualized in Figure 1.

These basic steps map to the following commands:

- Get Local Copy: `svn checkout`
- Make Changes: `svn {add,move,copy}`
- Review Changes: `svn {status, diff}`
- Fix Mistakes: `svn revert`
- Get Changes: `svn update`

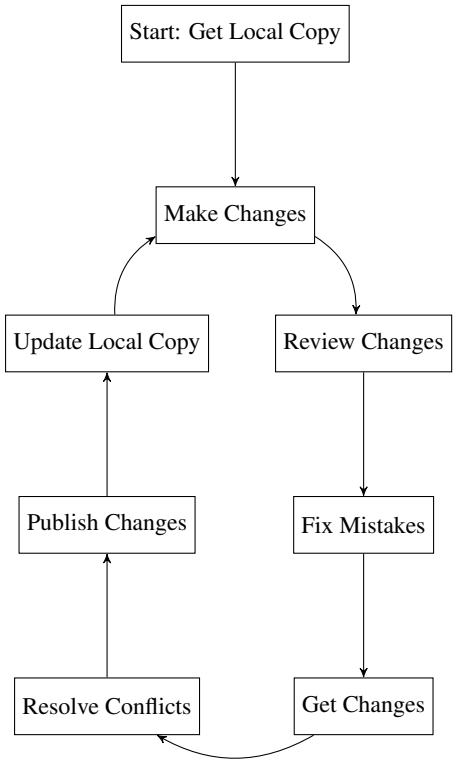


Figure 1: Basic Workflow in SVN

- Resolve Conflicts: `svn resolve`
- Publish Changes: `svn commit`
- Update Local Copy: `svn update`

Note that when resolving conflicts, a developer must first attempt to update their local copy with an `svn update`.

2.4 Distributed VCS

A Distributed VCS (DVCS) works on the principle that there does not need to be a central, global repository. Instead, it provides the flexibility to adopt different development styles by allowing multiple upstream repositories to be used. Git, a highly popular DVCS, is built around the concept that every copy of the repository is actually a repository that others can clone and submit changes to. This enables workflows such as the one adopted by Linus Torvalds (the creator of Git) in the development of the Linux Kernel: the Benevolent Dictator [2]. In this model, the dictator (Torvalds) maintains a *blessed repository*: a reference repository that is the "official" version. Developers update their local repositories from the blessed repository but publish changes to the dictator's trusted lieutenants. These lieutenants collect and merge these changes, providing a level of quality control and acting as an aggregation layer before code gets to the dictator. The dictator then merges changes provided by the lieutenants and publishes their master to the blessed

repository.

An interesting aspect of this style of version control is that it does not preclude the possibility of using a centralized workflow; as such, Git hosting systems like GitHub or BitBucket are frequently used as a central repository. This has one major implication: the visual language designed for a DVCS must be a superset of the visual language for a CVCS. A sample Git workflow is shown in Figure 2.

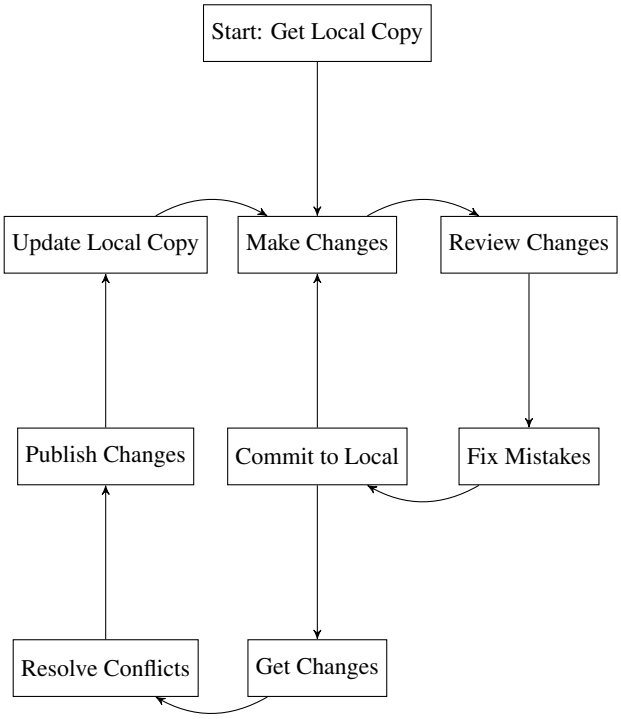


Figure 2: Basic Workflow in Git

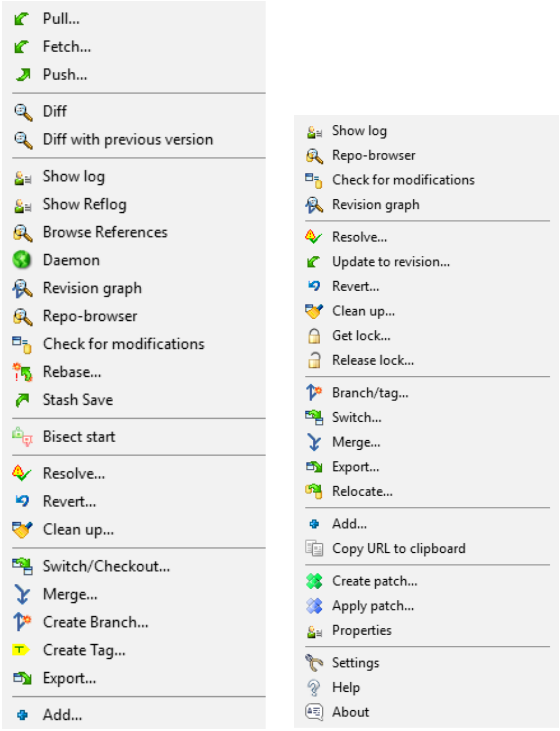
These basic steps map to the following commands:

1. Get Local Copy: `git clone`
2. Make Changes: `git {add,mv,rm}`
3. Review Changes: `git {status, diff}`
4. Fix Mistakes: `git reset`
5. Commit to Local: `git commit`
6. Get Changes: `git pull`
7. Resolve Conflicts: `git add`, `git commit`
8. Publish Changes: `git push`
9. Update Local Copy: `git pull`

Note that when resolving conflicts, a developer must first attempt to update their local copy with a `git pull`.

2.5 IDE Integration and VCS Clients

As version control is such an important tool in software development, multiple graphical interfaces for Git and SVN exist, and several IDEs have VCS Integration. Some IDEs, such as Netbeans, do not use icons in a toolbar and instead rely on drill down menus that do little to hide unneeded complexity. Figure 3 shows the icons that are used in the popular TortoiseGit and TortoiseSVN integration tools. These two programs seem to try to keep the wording and icons somewhat consistent between the two VCS.



(a) TortoiseGit context menu (b) TortoiseSVN context menu

Figure 3: Various VCS clients and IDE VCS Integration.

3 Designing a Visual Language for VCS

3.1 Design Principles

The visual language for VCS (VeCVL) consists of two sets of icons: those for DVCS (dVeCVL) and those for CVCS (cVeCVL). VeCVL follows these design principles (listed in order of importance).

3.1.1 Centralized Version Control is a Subset of Distributed Version Control

Any icon in cVeCVL must have meaning in dVeCVL. At this point in time, VeCVL is equal to dVeCVL; however, if new paradigms of VCS are introduced, the language will be able to grow to accommodate these advances.

3.1.2 Related Operations Have Related Icons

Due to the nature of the local and remote architecture, some operations have related meaning. For example: updating the local copy makes the developer's local repository match the remote repository. Publishing modified code makes the remote repository match the developer's local repository. It follows that the icons for these operations should be similar, but in some way opposite.

3.1.3 Aggregate Operations Have Aggregate Icons

Some operations are aggregates of more primitive operations. The `git pull` command, for example, can be achieved by performing a `git fetch` (that pulls indexes of remote changes) and then running `git merge` to merge the remote changes into the local repository. If possible, the icon for `git pull` should be the composition of `git fetch` and `git merge`. This should never supersede meaning across VCS, however; if such an aggregate icon would not make sense for both SVN and Git (and other VCS), then this design principle can and should be ignored.

3.2 Verb Meanings

To design a meaningful visual language, the operations must be grouped by their basic functionality. Note that in Figures 1 and 2, the workflows are presented independently of the verb chosen by each VCS. This is important because between the two systems, the same verb may have the different meanings for Git and SVN (`commit` being the most obvious example). As such, the first step in creating a visual language is to identify an independent set of verbs that communicates the meaning of the operations. These verbs should, when possible, not be biased towards a specific VCS. Table 1 shows the identified verbs and their meanings in VeCVL. The definitions use the following glossary:

- **Upstream**: A remote repository; in a CVCS, the central repository.
- **Local**: The local working copy of an upstream repository.
- **Conflicts**: a portion of source code that has been modified both locally and upstream.

3.3 Design of Individual Elements

Because aggregate operations should have aggregate icons, there should be some commonality between elements of the language. To this end, individual parts of meanings have graphical cues that combine to form the icons.

Repositories themselves will be represented as a directory (folder); they are stored on the file system as such, so the imagery associated with it will closely mirror what it represents. Upstream repository storage will be represented by a cloud, as cloud computing has become a catch-all term to describe services and platforms running on remote hardware not under control of the user. This concept of storing shared data in the cloud should be familiar to both veteran developers and students of software engineering and provides a simple graphical element that can be used in conjunction with others to give meaning. In contrast, local storage will be represented as a computer (similar to My Computer in Microsoft Windows).

Motion between repositories or files will be represented with arrows, with the direction of the arrow indicating the target of the action. Arrows are particularly important because they are capable of representing a merge operation in addition to showing the direction of motion. The idea of a change will be represented by the Greek letter Delta (Δ), which is common notation. Other common visual cues will be drawn from established software and icons. Figure 4 shows some of the design elements that comprise VeCVL.

3.4 Resolving Conflicts In Resolving Conflicts

Git and SVN handle resolving conflicts in files in different ways. In an SVN repository, when a conflict is detected, multiple versions of the conflicted file are created. The developer then populates the file that will be kept, and runs the `svn resolve` command, indicating to the VCS that the version has the desired content.

Because Git handles the local working copy as a fully qualified repository, the normal behavior for handling merge conflicts is different. The file in question has indicators inserted that show where the conflicts occur. The developer makes changes to this file, ensures the content is correct, then simply `add`s the file to the list of changes, and makes a local commit. This `git add; git commit` workflow has the same purpose as `svn resolve`: indicate to the VCS that the contents of the conflicted file are correct. However, it treats the changes like any other modifications to the file.

Because SVN creates a new command for this operation, VeCVL will also have a verb for this action (see Table 1, under `reconcile`). When mapping to commands, it maps directly to a `git add` and a `git commit` operation. To follow the design principles specified above, the icon should be an aggregation of the icons for add and commit. This combining simple operations into more complex operations is common in Git, as it is built into the very design principles of the system.

3.5 Fetching Changes Vs. Pulling Changes

Git has an operation designed to download a listing of modifications from an upstream repository *without* merging those changes into the local repository. This `git fetch`

Table 1: VeCVL Verbs

		VeCVL Verb	VCS Verb	Meaning
dVeCVL	cVeCVL	duplicate	git clone	Create a local copy of an upstream repository
			svn checkout	
		add	git add	Add a file to the list of changes
			svn add	
		delete	git rm	Remove a file from the repository and the file system
			svn rm,del	
		undo	git reset	Undo changes made by a add operation (remove file from list of changes)
			svn revert	
		move	git mv	Move a file from one location to another (renames the file)
			svn move	
		status	git status	View status of files in repository
			svn status	
		compare	git diff	Compare two files, or versions of same file
			svn diff	
		merge	git merge	Merge two conflicting versions of code (upstream and local, or multiple branches)
			svn merge	
		reconcile	git add,commit	Indicate that conflicting versions of a file have been handled
			svn resolve	
		sync	git pull	If no conflicts, make the local repository identical to the upstream repository
			svn update	
		publish	git push	If no conflicts, make the upstream repository identical to the local repository
			svn commit	
		commit	git commit	Save a commit in the local repository (local publish)
		fetch	git fetch	Get index of changes from upstream, but do not merge into local branch

operation allows a developer to pull in changes from multiple remote systems and use changes from all of them (by treating them as separate branches of code). This does in turn mean that `git pull` is in fact an aggregate operation: `fetch` the changes from a remote repository, and then `merge` the changes from the branch representing those changes.

This is not fundamentally different from the way SVN handles the `update` operation; it simply breaks the one complex operation into two more primitive operations. The final result still involves getting the changes from the server and combining them with the local working copy. Therefore, the icon for `sync` (`git pull` and `svn update`) can be a combination of the icons for fetching and merging, without destroying the meaning of the image.

(a) Repository (b) Upstream (c) Local (d) File

Figure 4: Basic Design Elements. Icons are Modifications of [7]

4 VeCVLv1

4.1 Language Specification

Figure 5 is the result of applying the design principles and elements enumerated in the previous section. Prototype icons express the idea of the visual language. Later iterations will include increasing readability by making better use of the icon space available.

Figures 6 and 7 show the same basic workflow diagrams shown in Figures 1 and 2 respectively, but use the visual language instead of the descriptions of the steps. It is easy to see that the workflow for SVN is just a subset of the workflow for Git, which was one of the design goals for VeCVL.

4.2 Implementing VeCVL

GitSubmit [1] describes the implementation of a simplified Git client used specifically for submissions in lower level courses. It was designed to simplify the use of interacting with version control systems and get students used to the basic workflow using Git. The GitSubmit interface implements the language described by VeCVL. A thorough description of GitSubmit, including screenshots, can be found in [1].

The UI for GitSubmit is designed around the principles of design for VeCVL; indeed, the UI itself is an implementation of VeCVL. Arrows are used to denote motion of changes and commits through the system.

5 Preliminary Results

Evaluation of VeCVL as implemented in GitSubmit is ongoing through user testing. The interface is being used as the exclusive submission system in two classes at Northwest Missouri State University: Data Structures and Algorithms.

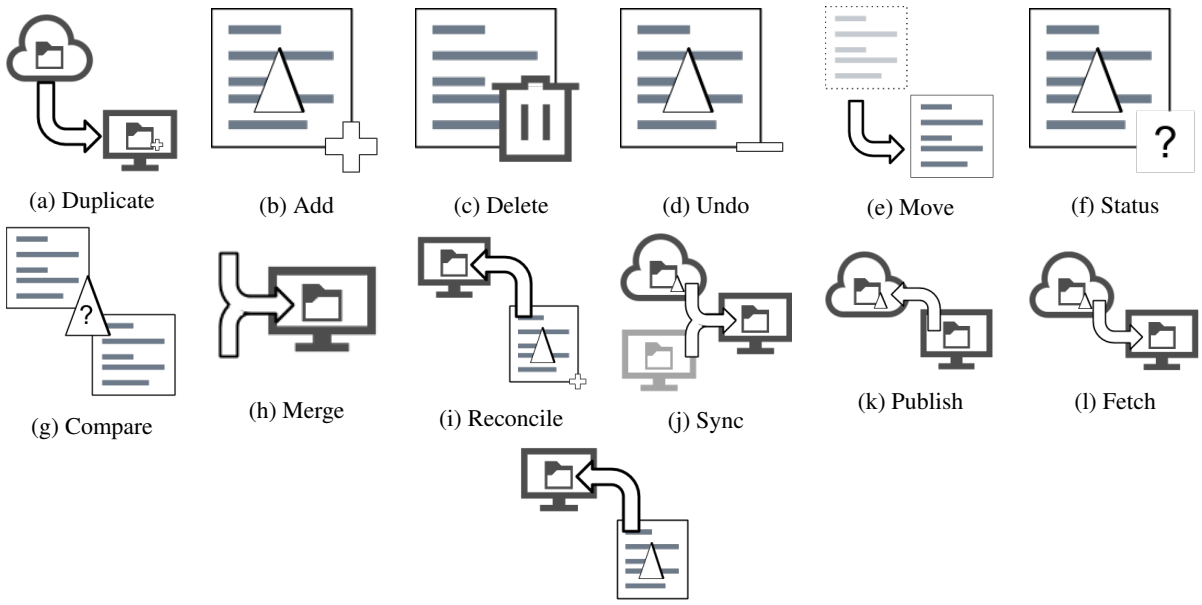


Figure 5: Proposed Visual Language for Basic VCS Operations

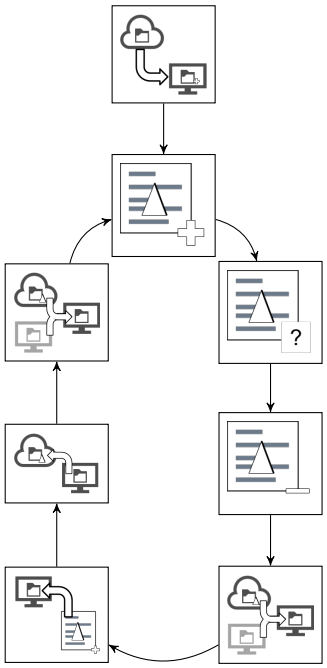


Figure 6: Basic Workflow in SVN using VeCVL

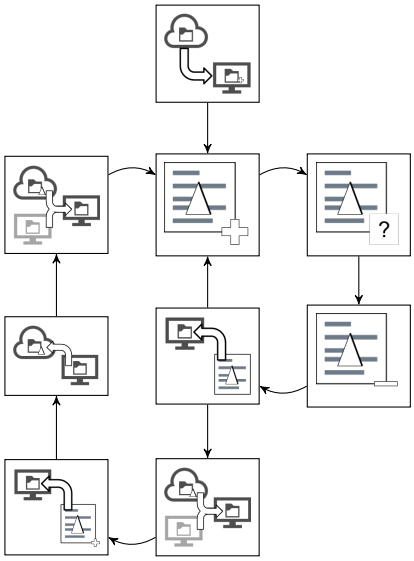


Figure 7: Basic Workflow in Git using VeCVL

Students were asked to fill out a voluntary survey about the process used in GitSubmit, and 32 responses were collected. This survey focused on certain Cognitive Dimensions of Notation. Two of the statements (with a Likert scale) directly relate to Error-proneness:

- The user interface helped avoid mistakes when submitting.
- The user interface made it easy to make mistakes when submitting.

The preliminary results for these questions is promising: 22 of 32 students either agreed or strongly agreed that the user interface helped avoid mistakes when submitting, while 8 were neutral and 2 disagreed. 20 of 32 students strongly disagreed or disagreed that the interface made it easy to make mistakes, while 6 students were neutral, and 6 students agreed with the statement. Additional feedback was collected that will be used to improve both VeCVL and GitSubmit.

6 Conclusions

Version control systems are important tools for developers working in any collaborative environment. Unfortunately, they are frequently introduced too late in a student's career for them to truly become habitual. Additionally, the complexity of interacting with these systems can scare students away from truly accepting them as a valuable part of their code development habits.

This paper introduces a visual language for version control systems as part of a multi-faceted approach to integrate version control into pedagogy. Together with [1], VeCVL aims to be a tool for teaching students how to use these tools easily. Additionally, it allows educators to teach students the *principles* of version control, not just "how to use Git" or "how to use SVN."

This initial version of VeCVL shows that it is possible to abstract away from system specific semantics and focus more on the generalized concepts behind version control. The visual language introduced in this paper covers the most basic tasks in a VCS, the ones that developers are most likely to use every day. It also follows defined design principles that should keep the language consistent and usable throughout its development. Preliminary user testing on a prototype implementation (GitSubmit) shows that the language can help with correctly using a Version Control System.

7 Future Work

This paper introduces the design and prototype implementation for a visual language for version control. The initial implementation covers basic concepts that developers use when working alone and covers the associated concepts across all of the major VCS. Future iterations are planned that will extend the visual language to include additional and powerful parts of VCS that developers use on a regular basis when working collaboratively.

Branching is a frequently used mechanism in VCS that allows multiple features and fixes to be developed concurrently without interfering with each other. The branching verb will be the next element added to this visual language. To adhere to the design principles expressed in this paper, other icons (specifically the merge/resolve action) may need to be revisited.

User testing is an important part of any visual language design. As the language is fleshed out, extensive user testing is planned that will encompass users of all levels of experience with VCS. Additional user testing will be done via surveys sent to developers in industry, academia, and students with varying levels of experience with these systems. The survey will attempt to target users of as many different version control systems as possible.

References

[1] D. M. Case, N. W. Eloe, and J. L. Leopold. Scaffolding Version Control into the Computer Science Curriculum. In *Proceedings of the 2016 International Workshop on Distance Education Technology (in conjunction with the 22nd International Conference on Distributed Multimedia Systems (DMS'16))*, 2016. In Preparation.

[2] S. Chacon. *Pro Git*. Apress, Berkely, CA, USA, 2nd edition, 2014.

[3] V. Driessen. A successful Git branching model, 5 Jan. 2010. http://nvie.com/posts/a-successful-git-branching-model.

[4] M. Guzdial. Software-realized Scaffolding to Facilitate Programming for Science Learning. *Interactive Learning Environments*, 4(1):001–044, 1994.

[5] J. Heer, M. Bostock, and V. Ogievetsky. A tour through the visualization zoo. *Commun. Acm*, 53(6):59–67, 2010.

[6] C. Jones, R. Armstrong, and K.-L. Ma. Visualizing the Commonalities Between Hierarchically Structured Data Queries. In *Proceedings of the 16th International Conference on Distributed Multimedia Systems (DMS'10)*, pages 251–256, 2010.

[7] KDE. GitHub - KDE/breeze-icons: Breeze icon theme. https://github.com/KDE/breeze-icons.

[8] M. Pilato. *Version Control With Subversion*. O'Reilly & Associates, Inc., Sebastopol, CA, USA, 2004. Accessed from http://svnbook.red-bean.com/en/1.7/svn-book.pdf.

[9] N. B. Ruparelia. The History of Version Control. *ACM SIGSOFT Software Engineering Notes*, 35(1):5–9, 2010.

[10] R. Wilson and B. Levine. Medical Gas Alarm System, Oct. 29 2015. US Patent 20,150,310,718.

[11] D. Wood, J. S. Bruner, and G. Ross. The Role of Tutoring in Problem Solving. *Journal of Child Psychology and Psychiatry*, 17(2):89–100, 1976.

A Browser-based IDE for the MUzECS Platform

Omokolade Hunpatin Casey O'Hare Ryan Thomas Dennis Brylow
Marquette University
MSCS Department – Cudahy Hall
1313 W. Wisconsin Ave.
Milwaukee, WI 53233
firstname.lastname@marquette.edu

***Abstract*—We report on a scalable, portable, and secure visual development environment for programming embedded Arduino platforms with Chromebooks in a successful secondary school computer science curriculum. Our web-based environment is part of the larger MUzECS project, an inexpensive replacement module for the Exploring Computer Science (ECS) course being widely deployed in United States high schools. Students use MUzECS to gain a deeper understanding of computing, through a set of blocks which provide appropriate abstractions for working with low-level hardware.**

MUzECS improves upon the existing curriculum module by reducing the hardware cost by an order of magnitude, while still preserving the key ECS pillars of computer science content, student inquiry and classroom equity. Programming with visual blocks provides a more attractive tool for introductory courses than traditional approaches, and yet enables high-impact exploration activities such as building a series of embedded musical instruments.

The current work combines and modifies several existing tools to eliminate technical barriers on low-cost platforms like Chromebooks, such as the reliance on special block-based toolchains, remote compilation servers, or multi-stage transfers for student code.

1. Introduction

According to the US Bureau of Labor Statistics, nearly 500,000 new jobs will be created in computing over the course of the next 10 years [14]. Computer Science drives innovation throughout much of the world economy, but it remains marginalized throughout primary and secondary school in many countries. Exploring Computer Science (ECS) [5] is a secondary school course which is currently being adopted in many parts of the United States. ECS was designed from the outset to address persistent gaps in representation by women and minority ethnic groups observed in the computing field [10]. It is targeted to early high school students (ages 14-16), and is designed to work well in traditionally under-resourced schools. However, the Achilles' Heel of ECS has proven to be its sixth and final curriculum module, which has, until the latest revision, relied on engaging, but costly and proprietary, robotics kits.

The MUzECS Project [2] was launched to provide a low-cost alternative to the ECS robotics module, and has been field-tested by hundreds of students in a dozen different school classrooms since its deployment. The physical platform which MUzECS operates on is a combination of an Arduino *Leonardo* or an Arduino *Uno* - inexpensive, commercially available credit-card sized embedded computers - and a "shield" - a circuit board extension that plugs into the top of the Arduino board to provide additional peripheral hardware. The Arduino is commercially available, but we produce the MUzECS shields in-house and provide them at cost. In contrast with current offerings for ECS module 6, our platform is open-source and can be easily extended to work with a variety of specializations within secondary school computer science, all at a very low price.

Recent trends in educational technology have led to a growing number of schools investing in Chromebooks [8], thin client laptops that run a Linux variant and the Chrome internet browser. For schools, Chromebooks represent inexpensive machines with lower maintenance costs and few of the device driver, application compatibility and software virus problems inherent in other types of personal computer. For computer science educators, Chromebooks encourage reliance on cloud-based services, but present new technical barriers to installing traditional software development tools, such as integrated development environments (IDEs), compilers, and debuggers.

The first release of the MUzECS programming dialect leveraged the prior Ardublock system, which could operate on any platform capable of running Java applications. Chromebooks, due to their very nature, do not allow such applications to run.

In this work, we present a powerful, web-based graphical programming environment for Arduinos and MUzECS shields, capable of running on stock Chromebooks as well as virtually any platform compatible with the Chrome web browser. Our solution consists of a browser-based IDE for Google Chrome, and a Chrome extension which allows for client-side execution of users programs. This is portable to more platforms than prior work, scales to a larger number of students with reduced load on webservers, and closes several usability and security issues with prior work.

DOI reference number: 10.18293/DMS2016-043

1.1. Initial Solution

In designing a block-based IDE for wide deployment in high schools, it is necessary to ensure that many students can use the tool simultaneously (scalability), and that they do not access other students' programs (uniqueness/security.) Finally, we must ensure that our system can be used on the wide variety of platforms used in high schools (portability).

MUzECS's first software platform was based on Ardublock, a Java-implemented graphical development environment which translates blocks directly to Arduino code. Ardublock is a stable add-on to the widely used Arduino IDE, and it handles the uniqueness and scalability problems effectively, as all code is executed on the user's own machine. Furthermore, since Java is a platform-independent programming language, this solution was quite versatile, capable of running on traditional Windows, Mac, and Linux OSes. Schools in the Milwaukee area, however, have been gradually adopting Chromebooks, a low-cost and lightweight laptop. Unfortunately, Chromebooks are only able to run the Chrome browser and Chrome Applications, and cannot install general purpose software like the Ardublock toolchain.

1.2. Arduino

The Arduino is a small microcontroller board with a universal serial bus (USB) plug to connect to your computer and a number of connection sockets that can be wired to external electionics such as motors, relays, light sensors, laser diodes, loud speakers, microphones, and more. They can either be powered through the USB connection from the computer, from a 9V battery, or from a power supply [12].

Our MUzECS shield attaches to the GPIO pins on the Arduino Uno or Leonardo. Our MUzECS shield consists of four LEDS, four buttons, a piezo speaker, and a distance ping sensor (Figure 1). We made blocks to manipulate all of the peripherals on our MUzECS shield.

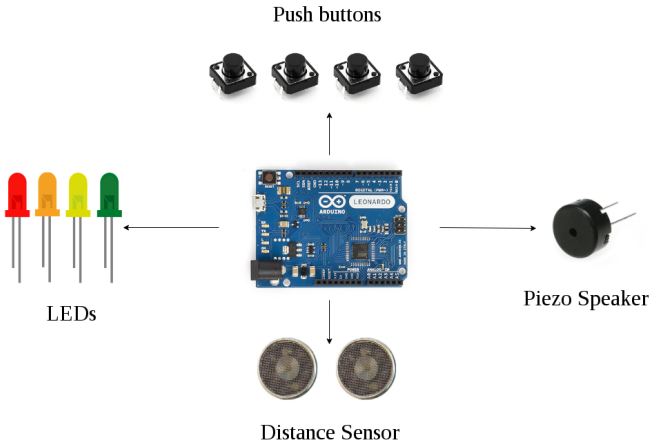


Figure 1. Peripherals for the MUzECS shield board

For a detailed breakdown of the capabilities of our different hardware platforms, see section 4.3: Hardware Support.

2. Related Work

Several platforms share the overall goals of the MUzECS project. Earlier versions of the ECS curriculum used LEGO Mindstorms robotics kits, which while highly versatile, were prohibitively expensive for many schools implementing ECS. Version 7 of the ECS curriculum will use more cost effective Edison Robots[3]. Like MUzECS, the Edison Robotics platform has a visual language designer, EdWare, where users assemble and connect blocks in a graphical environment to program. The Edison Robots themselves are LEGO robotics kits which are designed to move around. Like MUzECS' software, EdWare is free, and the Edison Robotics hardware is substantially cheaper than the LEGO mindstorms, ranging from $33-$50 per kit[3].

Code.org's CS Discoveries course is expected to make use of the Adafruit Circuit Playground[1], an all-in-one Arduino platform with similar peripherals to the MUzECS hardware, plus several attractive improvements, for an astonishing price below $20 per kit. At this writing, the Circuit Playground hardware is not yet widely available, (and the proprietary design cannot be replicated by third parties,) the full accompanying curriculum has not yet been released, the block-based tools for programming are progressing through beta versions, and pilot teachers have not yet received training for deploying the curriculum next spring.

Both the Edison and Circuit Playground alternatives were made available late enough in 2016 to miss the teacher training and deployment windows in the spring. The earliest opportunity for head-to-head comparison of these curricula in real ECS classrooms will thus be in spring 2017. In contrast, our work has been publically available for ECS teacher use since spring of 2015, and now will also be available on Chromebooks for the coming school year.

With MUzECS, Edison, and Circuit Playground all available for less than $60 per kit, the primary criteria for teachers to differentiate upon is the power and flexibility of the platforms. To that end, while MUzECS is the most expensive of the three, it is also the most open-ended with the clearest path forward for subsequent high school courses that would teach more advanced concepts, transition to text-based programming langauges, or be extendible with new shield hardware peripherals.

OzoBlockly [4] is another web IDE based upon Blockly. The OzoBlockly IDE pairs with the proprietary "Ozobot Bit," a small robot, which can move atop a surface using autonomous LEDs, sounds and infrared proximity sensing. OzoBlockly has a unique and creative way to load a program to the OzoBot Bit. Instead of using the serial port on a computer, the OzoBot Bit can simply be placed on the screen of the computer and identify a sequence of flashing lights as a program. This allows for a very simple upload process to a device for the user and expands functionality to mobile devices that otherwise could not program external hardware. OzoBlockly focuses on small games to entice

younger users, while MUzECS approaches the user from a musical perspective. Additionally, the Ozobot Bit does not allow for direct human interaction, unlike our MUzECS shield.

3. MUzECS Blocks

Our MUzECS block dialect was made to be translated to code that would run on our Arduino with peripherals. Designing the blocks has proven to be a key challenge in crafting MUzECS. We wanted our blocks to be easy to use, visually appealing, and we wanted each to manipulate a single peripheral on the MUzECS shield. The MUzECS blocks were made not only with the goal of teaching core computer science concepts, but also with the goal of facilitating the natural transition from block to text-based programming.

3.1. Our blocks

We based our visual programming dialect on Google's open-source, visual-block based programming environment called Blockly [9]. Our block-based interface works as follows: users attach blocks together in an online interface, in the Google Chrome browser. All blocks which the users assemble must be placed inside a main control block which reads "program" and has two slots: "setup" and "loop". The "setup" slot is a place for the user to initialize different hardware components on their board. The "program" slot is where the user assembles the blocks for logic and instructions.

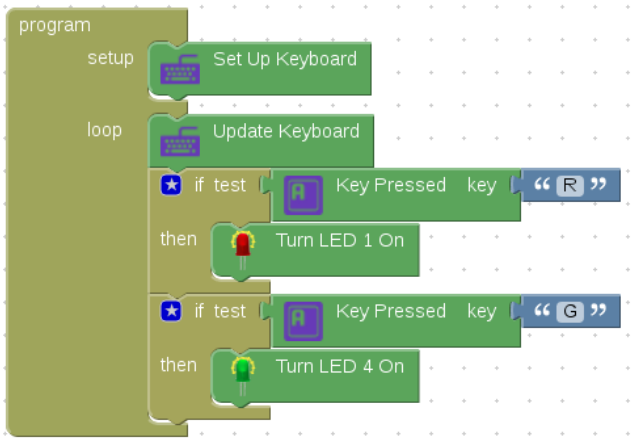


Figure 2. Example program which uses both the setup and loop slots of the program block

When approaching this problem, we had the option of doing all setup in the background, silently, or making it explicit. We made the decision to make setups explicit to help facilitate the transition to text-based languages, where initialization is a key concept. We should note that, when taking this approach, we chose to include simple error messages if the initialization is not performed. This stands in contrast to a programming language like Scratch, which does not include error messages. We go even further with the idea of initialization with the keyboard blocks, where "Set up Keyboard", "Update Keyboard", and "Key pressed" are all individual blocks. We chose to separate these actions each into their own functions rather than shadowing the behavior of the computer. We believe that these three separate blocks better help students understand the nuances of initialization and different objects. Another design decision that we made was to create drop-down menus for most of our operators, so that after a block has already been placed in the interface, it can be easily changed to a different operator without reconstructing the entire section.

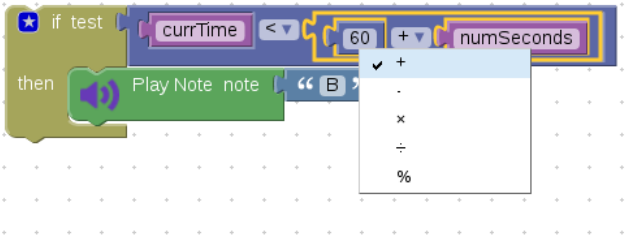


Figure 3. Clicking on the drop-down menu on the arithmetic operator lets the user select another operator

We have also designed our block-based interface to run in sequence, from top to bottom, and not to make our language event-driven. For more on this, see section 3.3 - Transition to Text-Based Languages.

We re-implemented the MUzECS dialect of Ardublock and some original Ardublock blocks to make the platform browser-based. Blockly's JavaScript API was used to create blocks, which already provided mechanisms for snapping blocks into place and generating code from blocks. Each drawer of blocks has a distinctive organizing color; the blocks that we implemented fall into several drawers: Control, ECS, Variables/Constants, Math operators, Logic operators, Communications, Advanced Pins and Advanced Code.

The "Control" drawer contains blocks which represent control structures commonly found in text-based programming languages: loops, if-then statements, and functions. Our goal with each of these blocks was to introduce students to control structures that they may use if they continue with programming. The "ECS" drawer contains the blocks for manipulating components of our Arduino hardware. We have blocks for: turning LEDs on and off, playing musical notes, playing notes for specific durations, reading the distance sensor, detecting button presses, reading host computer keyboard input, and reading the accelerometer values.

Blocks from the Control drawer and the ECS drawer are used in virtually every program that a student creates, so we paid extra attention to making these blocks carefully, and abstracting appropriately. The variables/constants drawer contain blocks that are used to create constants and variables to be used throughout the program. The math and logic operators drawers contain blocks to perform mathematical operations, get random numbers, and compare quantities. As

students progress and become more advanced programmers, they often use these blocks more frequently. Finally, the communications and advanced drawers contain blocks that can be used for debugging and more advanced operations - few students use these blocks, but we included them in case the instructors or especially advanced students would like to explore the interface on a lower level.

In the world of programming, there are always multiple ways to approach a problem. We wanted to assure that students using our block-based interface have multiple approaches available to them, so we built a diverse set of blocks into our interface. For instance, we have a "Play note" and a "Play note for a given time" block, and a "Play frequency" and "Play frequency for a given time" block - four different approaches to the same task. This kind of block diversity can be found throughout our interface. We aim for our interface to give students a minimal number of blocks, with a maximal amount of expressive power. We seek to strike a balance where students can express complex actions with simple blocks, and they don't have to labor over a complex block-based interface to perform relatively complex tasks. A prime example of this principle in action is the "Get Distance" block, which abstracts a substantial amount of JavaScript code into a single block.

3.2. Visualization of Functionality

Figure 4. MUzECS blocks with pictures that are associated with the function of the block.

Our top design criteria for the language is that it be simple to use for introductory high school students. Many blocks in MUzECS have pictures associated with them that are related to functionality. For example, "turn LED 1 on" block has a bright red LED picture on it, and "turn LED 4 on" block has a bright green LED picture on it. Likewise, "Button Pressed button #" has a picture of a hand pressing a button and "No Tone" block has a picture of a speaker with no sound waves coming out of it.

Blocks are generally colored according to functional category, such as green for I/O operations, and black for the Advanced Code block that allows users to inject raw text-based code into their program.

3.3. Transition To Text-Based Languages

Designing a system such that it facilitates a transfer of knowledge is an important, but not an easy task. Many MUzECS blocks were created not only because they are necessary to make our visual language complete, but also because they are similar to necessary control structures found in popular text-based languages (TBLs). For example, we have opted to make our programming language primarily structured around a single thread of control, despite the fact that JavaScript, our underlying language, lends itself more to an event-driven paradigm. This deliberate design decision was made because the dominant Arduino toolchains do not include support for multi-threading runtimes.

This decision was a crucial part of our design. It is noteworthy to contrast our model with another popular block-based language, Scratch, which aims to be an easier alternative in teaching people how to program. Scratch is developed at the Massachusetts Institute of Technology and is primarily event-driven. Scratch has been noted to fall short in facilitating the knowledge transfer to text-based languages, and its design as an event-driven language has been cited as a key reason why this is the case [2][7]. Another concept in TBLs which Scratch has struggled to establish is that of initialization [7]. We believe that using a single thread of control will help students to understand initialization, but we have also made an additional effort to teach initialization by creating setup blocks for certain hardware components, as mentioned earlier.

4. Chrome-based Arduino

4.1. Curriculum

Our curriculum is built on the existing secondary school Exploring Computer Science (ECS) curriculum. The ECS curriculum has proven to be a successful way to teach computer science to underrepresented groups in the past [11], and we specifically designed our curriculum in line with the goals and methods of ECS. The ECS curriculum abides by three guiding values: equity, inquiry, and CS content. ECS strives to be equitable by verifying that students from all backgrounds have a fair shot at learning about computer science. In practice, this means ECS teachers must choose assignments which all students have an (approximately) equal chance at understanding - and not choosing assignments that require students to understand the rules of chess, or giving extra credit on a test for a question about a sci-fi movie, for example. ECS also holds inquiry as a guiding value, meaning that, in classrooms, students and teachers should always be focused on discovering and asking questions, instead of constantly making assertions. Finally, ECS holds Computer Science content as an important tenant of the course. This may seem obvious, since the course is, after all, a computer science course. But its emphasis as only one of three guiding values is perhaps the most telling aspect of how ECS is meant to be taught. ECS is meant to be as much equitable and full of inquiry as it is about computer

science. This plays into the design of our system as well. We strive for our system not to just to teach computer science to students, but to be equally accessible for all students, and for it to instill a yearning for more computer science education.

We are confident that MUzECS fits the bill that ECS provides, which is exciting, given the past success of ECS. Simply put, we know our curriculum works, because we build it on the shoulders of a proven course.

4.2. A Scalable Platform

Our foremost concern when approaching this problem was developing a software system which was scalable. The central issue that we needed to solve was that of compiling the Arduino code; previously, we hosted the server which compiled the programs and sent the compiled code back to the users. We were then faced with a choice: on one hand, we could stick with the same model and throw more resources behind the compilation server. If we created a decent distributed system, we might have been able to outsource our compilation servers to a cloud-services-provider such as Microsoft or Amazon. In retrospect, this probably would have been a viable option. We reasoned, however, that we would be doing better if we could completely reinvent the model; it would be best if there was no server-side compilation at all. We know that networks and communication are inherently prone to eventual failure, so we think it best to remove them from the process, or minimize their role, if possible.

Our research revealed that it is, in fact, possible to program an Arduino without compiling Arduino code. If the Arduino is flashed with Standard Firmata [6] firmware, any client-side programming language which has an implementation of Firmata can be used to program an Arduino. Programming languages like Python, Perl, Ruby, JavaScript, Java, and more all have libraries for Firmata. We chose to use JavaScript, and more specifically, the Johnny-Five robotics framework, to program our Arduino.

Johnny-Five works by executing the JavaScript code directly on the host machine - the code doesn't compile down to Arduino code. The JavaScript code is executed using Node.js - a JS runtime which is specifically designed to build scalable network applications. As the program is being executed on the host machine, basic I/O instructions are transmitted to the Arduino board. With the Arduino Leonardo, the board is required to be connected to the computer while the program is running, and the instructions are sent to the Arduino via USB Serial. One should note that when we take this approach, the Arduino must be connected to a computer to run programs - we are unable to upload programs and run them from a battery. We have found, however, that, in the classroom, this point is more or less negligible - students almost always have their Arduino plugged in to their computer anyways. Finally, requiring the Arduino to be physically tethered does not hinder student development because the MUzECS shield does not have any moving parts.

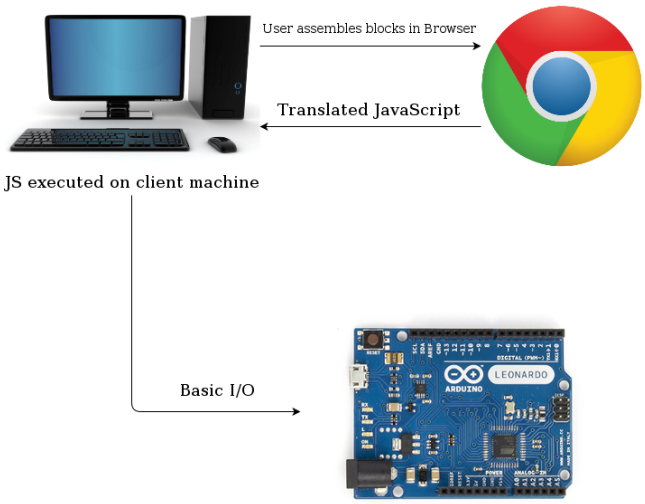


Figure 5. Design of our new system with Johnny-Five

Combining the different parts of this system was a relatively simple process. We used our existing block-based interface (Figure 4), which runs in the Chrome Application, and wrote the block definitions such that they translated directly to JavaScript code which satisfies the Johnny-Five framework. Our new architecture also works completely offline, once the application is downloaded.

Altogether, we believe our current software platform satisfies our original three considerations. It ensures uniqueness and security - we know that a student is only capable of running programs that they wrote on their own computer, on their own Arduino, because it must be physically connected.

Our platform is considerably more scalable than it ever has been. The burden of compiling programs has been shifted from the server to each user's computer. If a substantial influx of schools begin using our platform, we believe that our newly designed system will be able to manage the load with great efficiency. Finally, we believe our software platform is highly portable - perhaps as portable as is even possible for a modern software system. It is capable of running on any computing system that can download Google Chrome - meaning that our IDE can run on Chromebooks, Windows systems, Macs, and even the majority of Unix systems - again, virtually every computing system that is in use by modern high schools. By creating this software system, we believe we have combined the positive aspects of the previous versions of MUzECS and introduced new ones to make a secure, scalable, and portable platform.

4.3. Hardware Support

When the MUzECS platform was first launched, the only hardware that we supported was our own MUzECS shield (Figure 6) for the Arduino Leonardo.

Our shield was designed with a few considerations in mind; the foremost was to be cost-efficient. This is because we sought to design a cheap alternative to the most common

Figure 6. MUzECS shield

sixth module of ECS, which is incredibly expensive. Our design, uses a small set of peripherals - a speaker, four LEDs, four buttons, and a distance sensor. In order to avoid driving our cost up, we elected to build a third-party distance sensor into our design, which, consequently, does not have wide hardware or software support.

All of the hardware on our original MUzECS shield is still functional, even in our new model. We did, however, have to extend the firmata protocol in order to write JavaScript code for our distance sensor (refer to section 3.4 - Extension to Firmata Protocol for more). Still, the scope of what our shield hardware can do is somewhat limited. It should be noted that while the MUzECS package isn't the cheapest on the market, it has a wider range of hardware compatibility than other, similarly priced, options. The Edison Robot - a new installment in ECS curriculum v7 - is ten dollars cheaper than MUzECS, but it lacks the versatility of the Arduino Leonardo. Additionally, since the Edison Robot is a new addition to ECS, we have been unable to observe its effectiveness within the curriculum. We will be unable to present any sort of comparison between MUzECS' effectiveness and the Edison's effectiveness until Spring 2017, when teacher's use the platforms side-by-side in the classroom.

The Adafruit Circuit Playground (CP) is another option which has unique peripherals and is cheaper than our MUzECS package. Unlike the Arduino Leonardo, the CP (which is based on the Arduino *Flora*) has all of the necessary peripherals built-in. On the Circuit Playground, a circular board less than 2 inches in diameter, there are 10 RGB LEDs (all capable of emitting any RGB color), a piezo speaker, a triple-axis accelerometer, a light sensor, a sound sensor, a thermometer, and 8 capacitive touch pads which also act as general purpose input/output (GPIO) pins. The CP provides another low-cost alternative to more expensive robotics kits, and we plan to fully support the board with a set of blocks and curriculum in the near future.

4.4. Extension to Firmata Protocol

The open-source Firmata project provides a flexible protocol for remotely controlling a variety of embedded platforms from a tethered personal computer [13]. When an Arduino is imaged with the Standard Firmata sketch, a variety of programming language libraries can be used to issue platform-dependent instructions to the embedded hardware via a serial connections such as USB. The existing Firmata infrastructure was already suitable to manage most of the peripherals of the MUzECS hardware.

In this work, we have extended the Firmata protocol to support the MUzECS hardware ultrasonic distance sensors, a low-cost analog circuit not generally found on other platforms. Leveraging Firmata in our design has a number of advantages. The protocol is lightweight, requiring only 3 data payload bytes to set one or all the digital pins on the board [13]. Additionally, the Standard Firmata is the default firmware installed on every Arduino board, so it is readily available to students for upload to the Arduino. Our extended version of the Standard Firmata firmware, which we call MUzECS Standard Firmata, is similarly open-sourced, and included upload instructions for instructors and students.

Use of the Firmata protocol allows us to remove server-side compilation of the Arduino code, using the client local machine to run MUzECS block programs on the Arduino. The Standard Firmata is C++ Arduino code that runs on the Arduino and follows the client-server model, in which the Firmata sketch is the server running on the Arduino, and a client running on the user's computer issues commands to the embedded board.

Under the hood, the client sends *SysEx* messages to the Firmata server, which then executes actions on the Arduino. The standard command *SysEx* messages begin with a start byte (0xF0) and end with an end byte (0xF7). In between the start and end byte are 7-bit bytes which contain the commands one wants to send to the Firmata server. (See revelant segment of Firmata Protocol grammar in Figure 7.)

The Johnny-Five middleware doesn't directly support the MUzECS hardware ultrasonic distance sensor, but can pass extension commands through the existing Standard Firmata interface.

Our extension to the Firmata code base and protocol adds direct support for the MUzECS ultrasonic range finder. A new command type, *GetDistance*, was added to the existing *SysEx* command with byte 0x02. The client sends a *GetDistance SysEx* command to the MUzECS Standard Firmata firmware. After the firmware recieves the *GetDistance* command, it runs the code that activates the distance sensor. When the distance code completes, it sends the *GetDistance SysEx* command back to the client with the distance integer split in three 7-bit bytes. Among other low-level details, the Firmata code for handling the ultrasonic distance sensor performs smoothing of the data using a moving average, an important noise-cancelling step that allows block-based programs to produce useable musical input data.

<SysExMessage>	→	*<StartSysEx>* *<SysExCommand>* *<Data>** *<EndSysEx>*
<StartSysEx>	→	`0xF0`
<SysExCommand>	→	`0x00 - 0x7F`
<Data>	→	`0x00 - 0x7F`
<EndSysEx>	→	`0xF7`

Figure 7. A BNF grammar of the format of *SysEx* messages

5. Future Work

One of our biggest concerns that we sought to address throughout this entire project was making MUzECS modular and reusable. First, we created our software to be modular by being able to support multiple hardware platforms. With new boards come new opportunities to modify our curriculum and deliver it in new and exciting ways. Furthermore, our open-source collaborators on this project have developed ways to communicate with Arduino via bluetooth, wi-fi, and even raw TCP sockets. These are all exciting possible extensions that we could make to MUzECS in the future.

Beginning in the spring of 2016, we deployed MUzECS into several pilot high schools in the Milwaukee area. We have been collecting survey data regarding students' usage of blocks. In the future, we will need to analyze this survey data and make appropriate improvements to our platform, on both the hardware and software side. It may also be a good idea to build into the system a mechanism for collecting block usage data automatically as we scale it out even further.

6. Conclusion

MUzECS's original goal was to create a cost-effective, block-based platform for the sixth module of the Exploring Computer Science curriculum. This paper presents technical solutions that extend this work to an increasingly common device in school classrooms, the Chromebook. The resulting system retains the carefully tuned block-based programming environment that has been specially adapted to hardware and curriculum widely used to broaden participation in introductory computer science courses. The contributions of this new version include improved scalability and security over previous browser-based solutions, with the advantage of greater flexibility for expansion to more advanced coursework than several alternative systems that will be piloted in ECS classrooms in spring of 2017.

Prior work studying the effectiveness of student learners transitioning from block-based languages to text-based languages has identified shortcomings related to understanding of initialization. Our current system includes decisions in the block language deliberately designed to address these problems, smoothing the transition to text-based.

The next steps in this work are to complete data collection from actual ECS classrooms that are using MUzECS on Chromebook, and to compare usability and student learning with alternative platforms.

Acknowledgements

The MUzECS project is supported in part by the National Science Foundation, grants CNS-1339392 and ACI 1461264. Our thanks to Rick Waldron and the Johnny-Five contributors, and Luis Montes for his contributions to the Chrome implementation. The anonymous VLC reviewers provided extensive feedback that improved the final version of this paper.

References

[1] Adafruit. Circuit playground, 2016. https://www.adafruit.com/product/3000.

[2] M. Bajzek, H. Bort, O. Hunpatin, L. Mivshek, T. Much, C. O'Hare, and D. Brylow. Muzecs: Embedded blocks for exploring computer science. In *Blocks and Beyond Workshop (Blocks and Beyond), 2015 IEEE*, pages 127–132, Oct 2015.

[3] Edison Robotic. https://meetedison.com.

[4] Evollve, Inc. Ozobot and OzoBlockly, 2016. http://ozoblockly.com/.

[5] Exploring Computer Science. ECS v7.0, 2016. http://www.exploringcs.org/curriculum.

[6] Firmata Project. Firmata firmware for arduino, v2.5.3. https://github.com/firmata/arduino.

[7] D. Franklin, C. Hill, H. A. Dwyer, A. K. Hansen, A. Iveland, and D. B. Harlow. Initialization in scratch: Seeking knowledge transfer. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, SIGCSE '16, pages 217–222, New York, NY, USA, 2016. ACM.

[8] Google. Chromebook. http://www.google.com/chromebook/.

[9] Google Blockly. http://developers.google.com/blockly.

[10] J. Margolis. *Stuck in the Shallow End: Education, Race, and Computing*. The MIT Press, 2008.

[11] J. Margolis, J. Goode, and G. Chapman. An equity lens for scaling: A critical juncture for exploring computer science. *ACM Inroads*, 6(3):58–66, Aug. 2015.

[12] M. Simon. Programming arduino. In *Programming Arduino*, pages 7–8, Oct 2012.

[13] H.-C. Steiner. Firmata: Towards making microcontrollers act like extensions of the computer. In *NIME*, pages 125–130, 2009.

[14] United States Bureau of Labor Statistics. http://www.bls.gov/ooh/computer-and-information-technology/home.htm.

Social Network Models for the TDR System

YingJie Tang, HaoRan Zhang, ZhiJian Liang and Shi-Kuo Chang
Department of Computer Science
University of Pittsburgh, Pittsburgh, PA 15260, USA
yit20@pitt.edu, colinzhang@cs.pitt.edu, zhijianliang@163.com, chang@cs.pitt.edu

Abstract—The TDR system is an experimental multi-level slow intelligence system for personal health care. The TDR system can be used by a single user or a group of users who will interact to understand, maintain and improve each other's state of health. In this paper, we simulate social networks by applying the Abstract Machine model. Two social network models are described: the Circulated Model and the Teacher Student Model. We have incorporated the Teacher Student Model into the Chi super-component of the TDR system and obtained positive experimental results.

Keywords— Personal health care system, slow intelligence system, social network, abstract machine model.

1. Introduction

An experimental multi-level slow intelligence system for personal health care, called the TDR system, was developed as a test bed for exploring and integrating different applications in personal health care, emergency management and social networking [1]. The TDR system mainly consists of three super-components: Tian, Di and Ren. According to the Chinese philosophy these three super-components are the essential ingredients of a human-centric psycho-physical system. They can be thought of as human beings (Ren) interacting with the environment consisting of heaven (Tian) and earth (Di). For personal health care, there is a fourth higher level super-component called Chi (or Qi), which in this context represents the state of health of a person (or persons).

Decision making in TDR system is through multiple computation cycles involving the super-components to increase the chances of survival and well being of human beings (or groups). Any action based on only one aspect of the environment without considering the other aspects could reduce the chances of survival, thus iterative, multiple computation cycles are crucial for the TDR system.

The TDR system can be used by a single user, or a group of users who will interact to understand, maintain and improve each other's state of health. In this paper, we investigate the social network models for the TDR system.

The paper is organized as follows. Section 2 presents the GUI interface of the TDR system so that the reader can understand how the TDR system works in practice. In Section 3 an Abstract Machine model for the computation cycles is presented. In Sections 4 and 5 we show how and what part of the Abstract Machine model is extended to incorporate the social network model into interaction models. Based upon this approach, we describe two social network models: The Circulated Model (Section 6) and the Teacher Student Model (Section 7). We have incorporated the Teacher Student Model into Chi component in the TDR system and obtained positive experimental results (Section 8). In Section 9 we discuss the implications and further research.

DOI reference number: 10.18293/DMS2016-049

2. The Web GUI of the TDR System

In this section, we describe the Web GUI of the TDR system so that the reader can understand how the system works in practice.

The dashboard is the main GUI interface of the TDR system. As illustrated by Figure 1 it provides a high-level overview of the data in the system. On the left side, it has a menu panel that contains all the actions the user can perform, including activating and deactivating super-components. For the admin user, this menu will also include addition, deletion and modification of regular users.

Figure 1. The dashboard for Web GUI of TDR system.

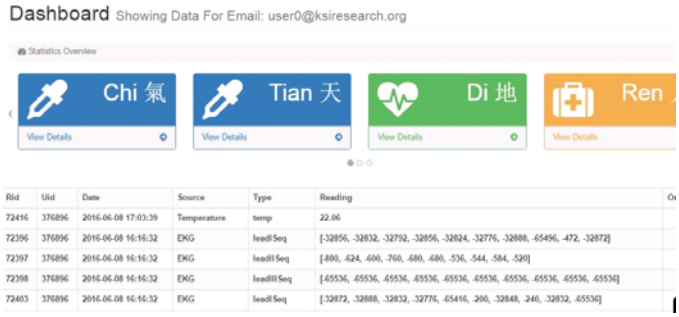


Figure 2. The Carousel.

There is a carousel that displays all super-components in rotation, four at a time for the PC screen and only one for the smart phone screen. This vividly demonstrates the idea of computation cycles in the TDR system. A component's banner is in *tranquil* state (blue or green color) until an alert is received and then it changes to *elevated* state (orange or red color).

As shown in Figure 2, when the user clicks on the "**View Details**" button at the lower right part of the dashboard, a table will appear beneath the carousel panel to display all records that belong to the current user. For each entry, it contains the date and time of a record, the sensor type, the data type, the actual reading of the data, and the originator. This scheme allows flexibility and scalability, as in the future there might be more and more sensors added to the TDR system. In Figure 2, the first record is the room temperature from the Earth (Di) super-component, and the other records are the EKG recordings from the Human (Ren) super-component.

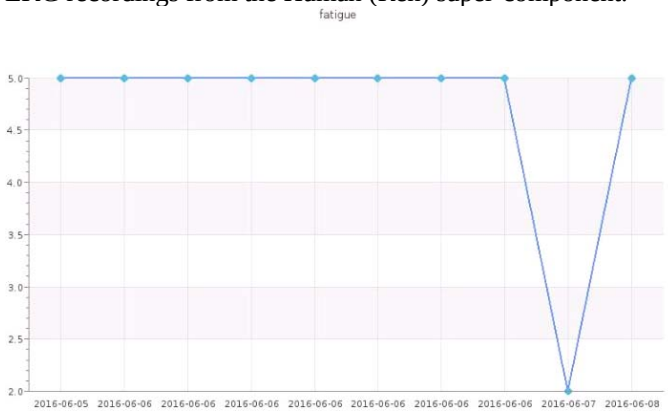


Figure 3. Visualization of fatigue in time.

If the user is communicating remotely with his/her doctor, a user might want to specify the record ID so that the doctor knows exactly what entry he/she is referring to. As shown in Figure 3, by clicking on the "**Draw Graphs**" button a graph showing the data-to-day changes of a selected data item can also be displayed by the GUI for visualization by the user or the doctor.

By clicking on the "**Analyze Data**" button at the lower right part of the dashboard, a user's records will be analyzed to evaluate his/her state of health such as the total-Chi of a person. Clicking on the "**Find Similar**" button will retrieve other user's records similar to the current user's record so that the user or the doctor can make a comparison to enhance his/her understanding. The details on data analysis and similarity retrieval are explained in the Appendix.

Interestingly, the user in a social group can use the same TDR GUI to retrieve other user's records, analyze them to evaluate their state of health, etc. If a particular user is a teacher or a 'master' with deep knowledge, this user can influence others to understand, maintain and improve one's state of health. Therefore the TDR system can also be used by social groups for collective personal health care.

3. The Abstract Machine Model for Computation Cycles

A slow intelligence system SIS typically possesses at least two decision cycles [2]. The first one, the *quick decision cycle*, provides an instantaneous response to environmental changes. The second one, the *slow decision cycle*, tries to follow the gradual changes in the environment and analyze the information acquired from the environments or peers or past experiences. The slow/quick decision cycles enable the SIS to both cope with the environment and meet long-term goals.

To model such decision cycles we introduce an Abstract Machine model of multiple computation cycles in Section 3.1, and then specify the computation cycles for the TDR system in Section 3.2. In Section 3.3 we describe multi-level computation cycles and how to apply the abstract machine model to social networks.

3.1. The Abstract Machine Model

The Abstract Machine Model is specified by: (P, S, P0, $Cycle^1$, ...,, $Cycle^n$), where

P is the non-empty problem set,
S is the non-empty solution set, which is a subset of Po,
P0 is the initial problem set, which is a subset of P,
$Cycle^1$, ...,, $Cycle^n$ are the computatin cycles.

Each computation cycle will start from an initial problem set and apply different operators such as $+adap_{Aij}=$, $-enum<$, $>elim-$, $=prop_{Aij}+$ and $>conc=$ successively to generate new problem sets from old problem sets until a non-empty solution set is found. If a non-empty solution set is found, the cycle is completed and later the same computation cycle can be repeated. If on the other hand no solution set is found, a different computation cycle is entered.

As an example the problem set P consists of problem elements p1, p2, p3, ..., p^n, and each problem element pj is specified by a vector consisting of attributes Aij. A computation cycle x will attempt to find a solution set by first adapting based upon input from the environment: $P^x0 +adap_{Aij}= P^x1$ that is to adapt based on attribute Aij, for example, by appending Aij to each element in P^x0 to form P^x1.

Then it may try to find related problem elements: $P^x1 -enum< P^x2$ where $P^x2 = \{ y: y$ is related to some x in P^x1, e.g. $d(x,y) < D\}$. Next it may try to eliminate the non-solution elements: $P^x2 >elim- P^x3$ where $P^x3 = \{x: x$ is in P^x2 and x is in $S\}$.

Finally the solution elements (or alert messages if there are no solutions) may be propagated to peers: $P^x3 =prop_{Aij}+ P^x4$ that is to export/propagate attribute Aij to peers.

Therefore this computation cycle can be specified succinctly as follows: Cycle^x [guard x,y]: P^x0 ***+adap$_{Aij}$***= P^x1 ***-enum***< P^x2 >***elim***- P^x3 =***prop$_{Aij}$***+ P^x4
The above expression is a specification of the computation cycle, not a mathematical equation. This expression should be read and interpreted from left to right.

If P^x4 is non-empty, the Abstract Machine will complete this cycle of computation and terminate at the end of Cycle^x, and it may later resume at the beginning of Cycle^x. Otherwise P^x4 is empty and the Abstract Machine will jump to a different Cycle^y. This is specified by [guard x,y] where x is the current computation cycle if a solution set is found (P^x4 is non-empty), and y is the computation cycle to enter if no solution set is found (P^x4 is empty). Before an Abstract Machine completes its current computation cycle, it will propagate the solution set (or alert messages) to its peers.

In the above, the elimination operator can be replaced by the concentration operator, whenever the solution set is not known apriori. The concentration operator applies a predefined threshold to filter out problem elements below the threshold: P^x1 >***conc***= P^x2 where P^x2 = {x: x is in P^x1 and th(x) above a predefined threshold t}.

3.2. Multiple Computation Cycles of TDR System

For the TDR system, a problem element is a combination of Tian, Di and Ren attributes. Those problem elements that are favorable for human survival are in the solution set S. The problem set P consists of problem elements p1, p2, p3, ..., p^n, and each problem element is specified by a vector consisting of the attributes from Tian (heaven), Di (earth) and Ren (human being), i.e.,

$$p^j = (t1j, t2j, ..., d1j, d2j, ..., r1j, r2j, ...)$$

For example, the Tian attributes tij are atmospheric variables such as amount of sunlight and water level, the Di attributes dij are residential variables such as ambient temperature and humidity, and the Ren attributes rij are personal health indicators such as blood pressure, EKG reading, heart rate, etc.

$$p^j = (sunlightj, waterlevelj, tempj, humidityj, bloodpressurej, spo2valuej, heartratej)$$

Initially some attributes may not be assigned any value and some may already have pre-assigned values. After most attributes have been assigned values one can decide whether the problem element is in the solution set. (The simplest case is that each attribute Aij has a solution range Rj, and if every attribute Aij falls within the solution range Rj then the problem element pj is in the solution set S).

In the TDR system, there are continuous interactions among the three super-components Tian, Di and Ren. Each super-component has its own computation cycle, which is basically the following: Starting from some problem set P0, the super-component first adapts to the input from the environment as well as from other peer super-components. It then tries to find related problem elements by enumeration. After those problem elements not in the solution set have been eliminated either using the elimination operator or using the concentration operator, the termination condition can be tested. The termination condition is expressed by [guard x, y] where Cycle x is the current cycle and Cycle y is the cycle to jump to. Whenever one super-component completes its computation cycle, if a solution is found the computation ends, otherwise the control is transferred to the next super-component. Since there are three super-components, we will have three computation cycles.

The Tian super-component has computation Cycle1:

Cycle1 [guard1,2]: P^10 ***+adap$_{Aij}$***= P^11 ***-enum***< P^12 >***elim***- P^13 =***prop$_{Aij}$***+ P^14

Likewise, the Di super-component has computation Cycle2:

Cycle2 [guard2,3]: P^20 ***+adap$_{Aij}$***= P^21 ***-enum***< P^22 >***elim***- P^23 =***prop$_{Aij}$***+ P^24

Finally, the Ren super-component has computation Cycle3:

Cycle3 [guard3,1]: P^30 ***+adap$_{Aij}$***= P^31 ***-enum***< P^32 >***elim***- P^33 =***prop$_{Aij}$***+ P^34

Notice the three computation cycles together form a higher-level computation cycle. High-level computation cycles are essential for a complex human-centric psycho-physical system such as the TDR system.

3.3. Multi-Level Computation Cycles

The three computation cycles together form a higher-level computation cycle. High-level computation cycles are essential for a complex human-centric psycho-physical system such as the TDR system. In the above specification, we can replace [guard1,2], [guard2,3] and [guard3,1] by [guardX,2], [guardX,3] and [guardX,1], respectively.

When computations in Cycle1, Cycle2 or Cycle3 is unsuccessful and solution set is empty, control is transferred to the next Cycle in cyclic order, i.e. first Cycle1, then Cycle2, then Cycle3 and then returning to Cycle1. On the other hand, when computations in Cycle1, Cycle2 or Cycle3 are successful and solution set is non-empty, control is transferred to the Chi super-component computation CycleX at the next higher level. CycleX is specified as follows:

CycleX [guardX,1]: P^X0 ***+$_{Aij}$adap***= P^X1 ***-enum***< P^X2 >***elim***- P^X3 =***prop$_{Aij}$***+ P^X4

In the above, the ***+$_{Aij}$adap***= may be the input propagated from the lower-level super-components, or from the super-components of other human observers (see below).

If computation in CycleX is unsuccessful, control is returned to the Tian computation Cycle1 (or the Di computation Cycle2, the Ren computation Cycle3, respectively). If computation in

CycleX is successful, then the computation terminates in CycleX and the Dashboard will display the results, i.e., the estimated total-Chi values.

The subjective evaluations can be entered by the principal user himself/herself based upon his/her subjective feelings. For example if he/she feels "sweaty at night", he/she will enter a value close to 10 (on a scale of 1 to 10) for the "sweaty-at-night" attribute for Chi.

It is also possible to formulate CycleX so that other human observers who are "friends" of the principal user can fill in the subjective Chi attributes This social network of human observers can also vote on updating the Chi attributes for the principal user. These human observers may even be allowed to fill in the objective Chi attributes as if they were sensors. Thus this TDR system is an iterative slow intelligence system, or what we call the *Sentient Net*.

4. Social Network Modeling

A social network can be modeled by a graph G = (P, A) where persons pi in the social network are modeled by nodes P and their relations are modeled by arcs A. A node can be attributed. For example, an attribute value 0.7 could mean a 0.7 probability of propagating its influence. An arc can also be attributed. For example, an attribute value 5 could mean 5 units of interaction between two persons pi and pk represented as nodes.

A person pi is I-related to pk if there is an arc between pi and pj and there are at least I units of interaction between them. A person pi is (I,D)-related to pk if there exists an I-related path from pi to pj and d(x,y) =< D is the shortest distance or minimum path length.

The slow intelligence operators are as follows:

Operator $enum_D$: P1 -enum< P2 where P2 = { y: y is (I,D)-related-to some x in P1, e.g. d(x,y) =< D}

Operator elim: P1 >elim- P2 where P2 = {x: x is in P1 and x is in S}

Operator $conc_t$: P1 >conc= P2 where P2 = {x: x is in P1 and th(x) >= t where t is a predefined threshold }

Operator adap: pj $+_{Aij}$adap= pk is to adjust pj based on input attribute Aij, for example, by appending Aij to pj with probability q, and not appending Aij with probability 1-q.

Operator prop: pj =$prop_{Aij}$+ pk is to output/propagate attribute Aij to pk with probability q.

We focus on the enumeration operator and with inputs of arguments, the output will be a system with super-components where enumeration will be performed in each cycle. In order to better formulate the problem, the following assumptions are made:

Assumption 1: All the elements in the problem set are connected to each other. That is, for every pair of elements x and y, d(x, y)<D is true.

Assumption 2: For every cycle in the system, each element having a probability less than 1 may be explored. In other words, not all the elements satisfying the condition d(x, y)<D will be explored.

Assumption 3: The solution set is achieved when the problem set is *stable* at cycle_t where cycle_i and cycle_j in the predefined range cycle_t-k to cycle_k satisfies the following condition: the difference between every attribute of a certain element in cycle_i and the attribute in cycle_j is smaller than a predefined threshold dij.

The first assumption simplifies the distance computation between elements. The second assumption restricts ourselves to real social networks. The third assumption is the definition for a stable solution for computation cycles.

We are formulating a social network by observing the influence of opinions. Thus picking a reasonable solution set is important. Opinions tend to be more steady when the profiles are fully interacted.

Our elements are P1, P2, … Pn, where each Pi represents a person in a group. Each person is defined by their [id, opinion, influencePercentage]. The id is a unique identifier for that person, and the opinion is the subject that would change during the computational cycles. Finally, the influence percentage is the indicator of how influential is that person (i.e. how capable is that person in changing others' opinions).

After knowing the nature of our elements, now we describe the problem set that contains those elements. Each problem set is divided into two sub sets: *High Influence Group*, and *Low Influence Group*.

The interaction happens in a sub-set basis. That is, H is interacting with L and vice versa. As expected, the result of this interaction is a potential change in the opinion of the element having interacted with the group. This change might happen, or it might not. It is a result of the average influence of all the elements in the interacting group.

Assuming that a change happened, there are two possible scenarios: the change in the opinion is either *influential* or *mind changing*.

The influential model is changing the element having interacted with group opinion based on the opinion of the

majority of the interacting group. For example, if the majority of the interacting group has the opinion X, then the change in element of the interacted group would be to the opinion X with some probability.

The other model causes the interacted group to change its opinion to the other one (in case of binary domain of opinions) regardless of what is the opinion of the interacting group.

Now we present the formal definitions of the model. From the abstract machine, we have the definition of the enumerator operator:

P1 -enum< P2 where P2 = { y: y is related-to some x in P1, e.g. d(x,y) < D}

This definition is extended in our model to the following:

P1 -enum< P2 where P2 = { y: y is x in P1, with a chance that its opinion is changed}

Each Pi is represented as a vector:

Pi = [id, opinion, influence], where id is a unique integer, opinion is a binary variable that represent an opinion, and influence is a percentage.

The set of elements is divided into two subsets (Figure 1):

H = {P1, P2, …, Pi}
L = {Pj, …, Pn}
Where H ∪ L = P

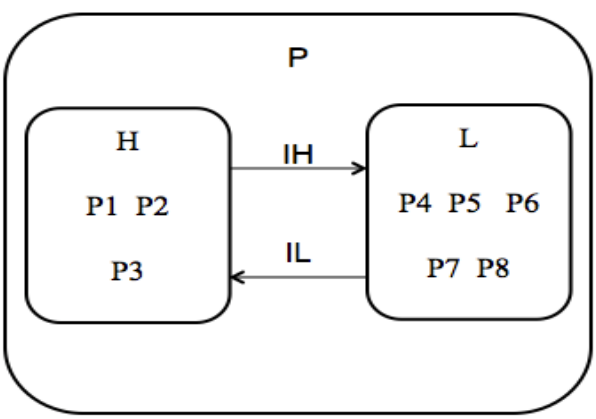


Figure 4. An example of a problem set with 8 elements.

During each interaction cycle, the interaction between the groups is represented by IH and IL, where IH is the percentage with which the group H affects the L group. The opposite holds for IL (see Figure 4). The probability with which an interaction will change an opinion of an element in the other group is calculated as follows:

$$IH = \frac{\sum_{i=1}^{|H|} P.influence}{|H|}$$

$$IL = \frac{\sum_{i=1}^{|L|} P.influence}{|L|}$$

Thus, each interaction cycle is a two-way interaction. From H to L (H → L) and from L to H (L → H). However, the percentage of changing the opinion of elements is different in each interaction.

5. The Interaction Models

When the the people element Pi is to have its opinion changed, there are two ways to change as mentioned above and discussed in [3]:

The Influencing Model: In this mode, if the majority of the interacting group has the opinion X, then the element that is interacted with will have the opinion changed to X (even if it is already has X as its opinion).

The Mind Changing Model: In this model, the opinion of the interacting group is irrelevant. Hence, if the element is going to change its opinion, it will change it to an opinion that is (not its current one) (e.g. flipping the opinion in binary domain opinions).

Some scenarios will now be presented. A video demonstration is available at: http://screencast.com/t/n48teU5c

We first present the influential model with two examples:

1) **Example 1**: | H | = 20, | L | = 180
H group influencing percentage = [7%, 10%]
L group influencing percentage = [1%, 3%]
5% in H have their opinion for '0'
95% in L have their opinion for '0'
Number of interactions is: 30 (bidirectional)
(See Figure 5).

2)
3) **Example 2**: | H | = 100, | L | = 100
H group influencing percentage = [80%, 90%]
L group influencing percentage = [10%, 20%]
5% in H have their opinion for '0'
95% in L have their opinion for '0'
Number of interactions is: 30 (bidirectional)
(See Figure 6).

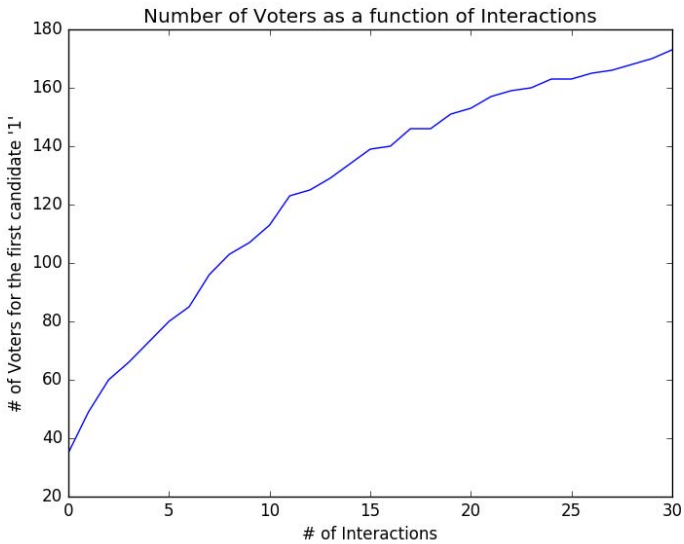


Figure 5. # of voters over # of interactions for Example 1.

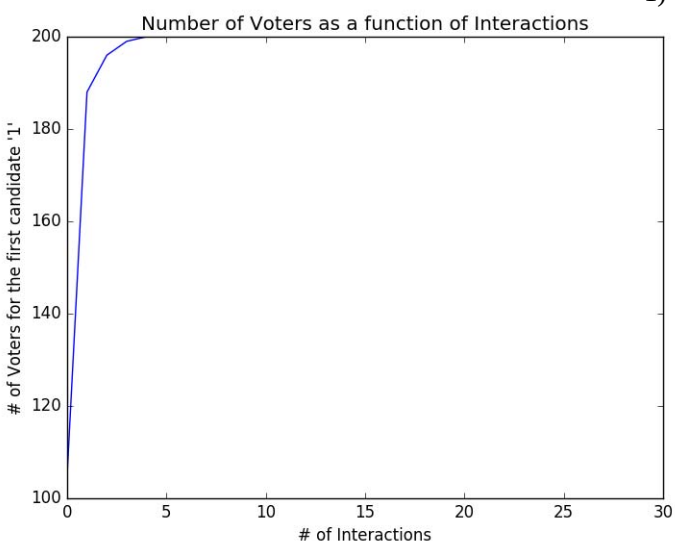


Figure 6. # of voters over # of interactions for Example 2.

From the above two examples, we see that the influence percentage controls the speed of convergence towards the "unified opinion".

Next we will show the behavior of the Mind Changing Model with the same two examples using the same parameters:

Example 3:| H | = 20, | L | = 180
H group influencing percentage = [7%, 10%]
L group influencing percentage = [1%, 3%]
5% in H have their opinion for '0'
95% in L have their opinion for '0'
Number of interactions is: 30 (bidirectional)
(see Figure 7)

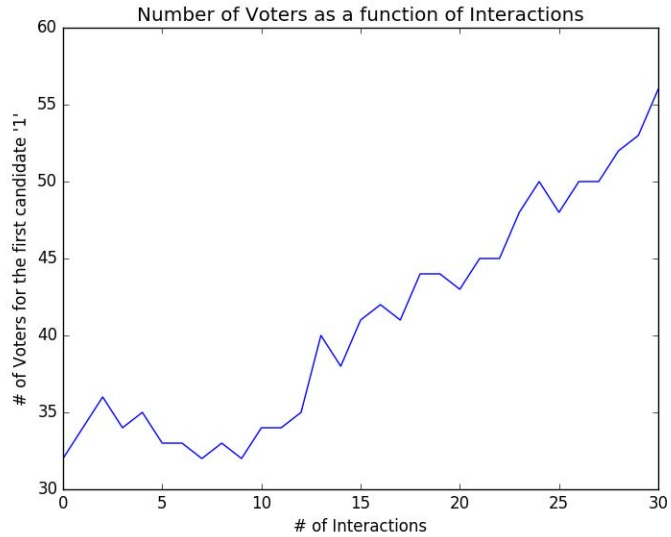


Figure 7. # of voters over # of interactions for Example 3.

1) **Example 4**:| H | = 100, | L | = 100
H group influencing percentage = [80%, 90%]
L group influencing percentage = [10%, 20%]
5% in H have their opinion for '0'
95% in L have their opinion for '0'
Number of interactions is: 30 (bidirectional)
(See Figure 8).

The interesting observation is that the behavior is not predictable. It changes due to people changing their minds, hence the fluctuation. The codes are available at this URL: https://dl.dropboxusercontent.com/u/19443460/code.zip

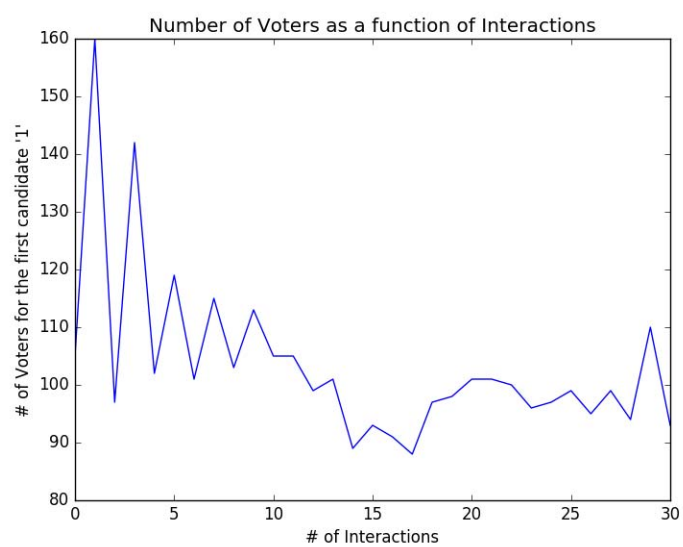


Figure 8. # of voters over # of interactions for Example 4.

6. Circulated Model

The Circulated model aims to weaken the Assumption 1 in the previous assumption in the social network model. The Assumption 1 made the social network model simpler by assuming that all nodes in the problem set are connected. However, this is nearly impossible in real social network models. Distance matrix is one of the most important features in the model and in this section, we introduce a distance matrix in our model to weaken the previous assumption.

In a circulated Model, communication is possible among all nodes albeit with different distance matrix. Adjacent nodes are connected with distance equal to 1 and the communication is bilateral, i.e. every node can influence its adjacent node with distance equal to 1. The influence is transitive, i.e. node1 can only influence node3 by passing the influence to node2 if node1 and node3 are not adjacent. A Circulated Model with n nodes is called Circulate-n model and a Circulate-5 model is shown in Figure 9.

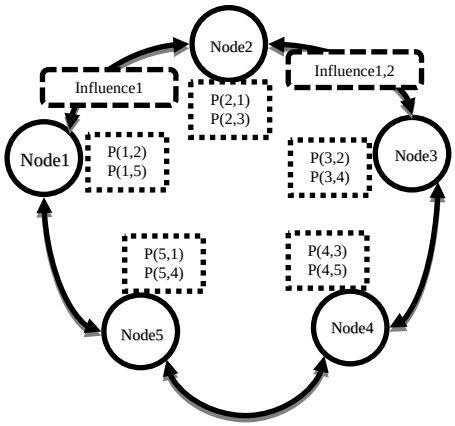


Figure 9. Circulate-5 model.

Node1 initiates an opinion influence and passes the influence to Node2, and Node2 passes the influence to Node3 by adding its' own views. In this case, Node1 influences Node3 by Node2 and it is one of the two ways that Node1 can influence Node3 whereas another ways is through:

Node1 -> Node5 -> Node4 -> Node3

Compared to real life social network models like Facebook or Twitter, the Circulated Social Network Model simplifies the relations of nodes and assumes that each node only interacts with its neighbors and the interaction to other nodes in the same group can be accomplished by interacting with its neighbors.

In this case, each node should store the probability of interacting with its neighbors. The distance matrix of each pair of nodes can be derived in the Circulate Social Network Model by the minimum distance of each path. In the above example, the distance between Node1 and Node3 is:

$$\text{dist}(\text{Node1}, \text{Node3}) = \min\{d(\text{Node1}, \text{Node2}, \text{Node3}), d(\text{Node1}, \text{Node5}, \text{Node4}, \text{Node3})\} = \min\{2, 3\} = 2$$

7. Teacher Student Model

Figure 10 illustrates not only the teacher student model but also how the TDR system actually works in social networking. Obviously the teacher student model fits the TDR system very well. In Figure 10 as well as the following scenario, message types are indicated by M9, M10, etc.

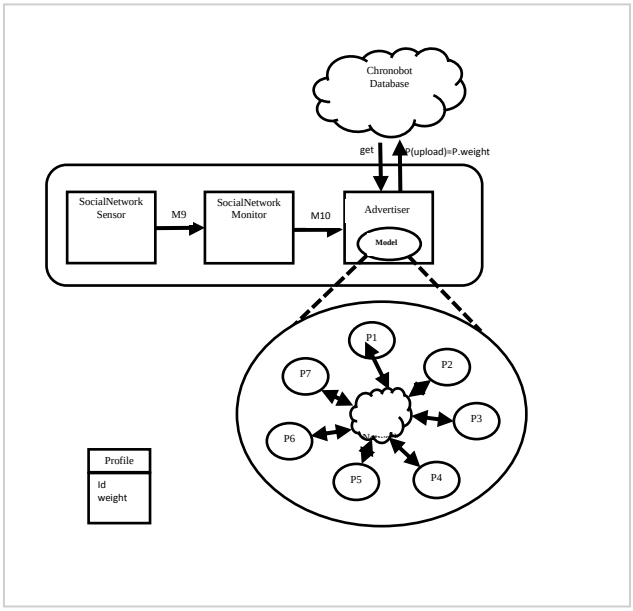


Figure 10. Teacher Student Model.

A scenario on how social network works in TDR system is as follows:

1. The SocialNetwork Sensor gathers information of Chi attributes (tongue, fatigue, weakBreath, pulse, sweaty) as well as the id and originator and send it to SocialNetwork Monitor through message M9.
2. SocialNetwork Monitor receives message M9 and calculates the total-Chi attribute using the same algorithm as from Chi attribute.
3. SocialNetwork Monitor sends the 6 attributes as message M10 to SocialNetwork Advertiser.
4. SocialNetwork receives message M10 and passes it to the SocialNetwork Model.
5. SocialNetwork Model gets the most recent value of the total-Chi attribute from Chronobot Database and compares it with the total-Chi from message M10 to evaluate the correctness of this record.
6. The SocialNetwork model adjusts the weights of each profile in the model and uploads the Chi attributes to Chronobot Database.

A profile in the group with higher weight is more reliable in judging others and we call this kind of profile *teacher profiles*. In the group model, there is one or more profile which is

called teacher profile who has a higher weight while the other profiles have lower weights. In order to propagate the weights to the Advisor component that uploads data to the Chronobot Database, we transfer the weight of each profile to the probability of uploading P(upload) = P.weight. In this case, whether a record of evaluations from a specific profile should be uploaded or not is associated with its weight attribute. The evaluation from a teacher is most likely to be uploaded to the database while an evaluation from a student is less likely to be uploaded. And an opinion from a student whose weight is small is less likely to be trusted and we upload this judgment to the database with a smaller probability. And thus all the records uploaded to the Chronobot Database are reliable. In this way, we transfer the weight of the profile to the trustworthiness of a judgment from a specific profile.

After a few computation cycles, students can learn the skills of n from the teacher and their weights will be increased so that they can produce more reliable evaluations. And it is the same as in the teacher-student relationship in the real life. As students learn from the teacher, the students will have gain the same power as the teacher. Once we make an assumption that the teacher's evaluations are always correct which means the teacher acts as an Oracle, the whole system will converge to a self-learning system and every profile in the system will be knowledgeable about the evaluation. And as the system improves, the evaluations will become more objective.

One important assumption we made in this model is that we restrict the number of profiles to be a small number, usually less than 10. A person is not likely to have a large number of friends that he can trust to evaluate his state of health. Since a small size group is easier to manipulate, this assumption makes the Teacher Student Model more robust.

8. Experimental Results for Teacher Student Model

In order to prove that the Teacher Student Model will converge to a stable status, we set up the following offline experiment. There are 7 members in the group including a teacher while the rest of the members are students. Teacher owns the weight of 1 which means every evaluation originated by the teacher is for sure reliable. On the other hand, students owns weights = 0.3 which means 3 in 10 evaluations originated by a student are correct and the rest are false. And with the hops forward, students can learn from the teacher and to gain the skill of judging correctly. If a student generates an evaluation that matches the teacher's evaluation, we increase his/her weights by a small value. And if he/she made a mistake, we decrease his/her weights in making evaluation. Thus only the weight changes in students worth tracing because the teacher is assumed to always generate correct evaluations. In other words in this Teacher Student Model, teacher acts as an Oracle and all the evaluations by the teacher are assumed to be correct.

Before starting the experiment, we create an *evaluations pool* where all the evaluations come from the teacher. The evaluations originated by students will be compared with those in the evaluations pool.

The experimental results show that as the hops (cycles) go on, the weights of each student increase steadily and all converge to 1 at the 14th hop. It successfully simulates the learning steps in real life. As the students learn from the teacher, they will gain the ability to evaluate state of health that the teacher has.

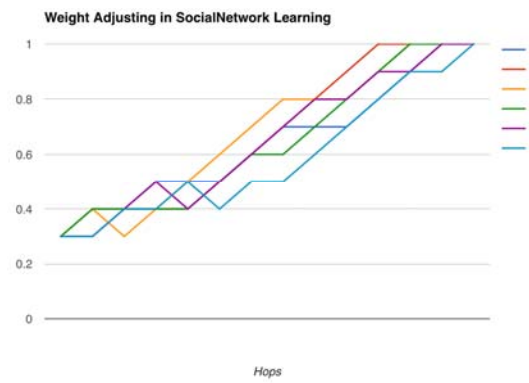


Figure 11. Experimental results.

In order to show how the learning converges, we did the following experiments: Three users, user0, user2 and user4 are incorporated in this experiment. User0 acts as the master in the group which means every made by user0 is set to be the standard. User2 act as a student and he should learn from the master how to make evaluations. Once user2 made an evaluation which is similar to the master, the reliability of user2 should be increased. Weight of user2 is the measurement of his/her reliability. Theoretically, as time goes by and user2 keeps on learning from the master, the weight of user2 should converge to 100 which is the same as the master. The last user incorporated in the experiment is user4, who is a patient. Both users0 and user2 make evaluations about user4. If the evaluation from user2 is similar to that of user0, we expect to see that the weight of user2 increases. The initial weight of user2 is set to 73. The master first made evaluations about user4 and the scores are (5, 5, 5, 5, 5), and the student later made an evaluation (5, 5, 5, 4, 5) about user4. We checked back the weight of user2, it increased to 74. This showed the ability of convergence of the Social Network Model in TDR system.

9. Discussion

One of our main goals is to expand the TDR system for the computation of Chi (also spelled as Qi in Chinese transliteration system HanYu PinYin). The Chi super-component is regarded as at a higher level. It has attributes including both objective measurements and subjective

evaluations [4]. This makes the Chi super-component both pro-active and adaptive at multiple levels.

The dashboard for TDR system can be further refined. When user clicks on “view details” for the Chi super-component, a list of attributes for Chi is shown. The objective measurements in this list is filled by the multi-level computation cycles based upon actual measurements. The subjective evaluations are entered by the principal user himself/herself based upon his/her subjective feelings, or his/her friends, teachers and a master with deep knowledge.

The convergence of the opinion shows that a steady opinion of the group is met which means the solution set is achieved under the three assumptions. The Teacher Student model successfully resolves the peer evaluation problem in TDR system and can play an important role in future generations of the TDR system. In the future, we plan to adjust the model to having more than one teacher, so that user can learn both from the master and from other teachers or selected friends.

References:

[1] Shi-Kuo Chang, JunHui Chen, Wei Gao and Qui Zhang , “TDR System - A Multi-Level Slow Intelligence System for Personal Health Care”, Proceedings of 2016 International Conference on Software Engineering and Knowledge Engineering, Hotel Sofitel, Redwood City, USA, July 1-3, 2016, 183-190.

[2] Shi-Kuo Chang, "A General Framework for Slow Intelligence Systems", *International Journal of Software Engineering and Knowledge Engineering*, Volume 20, Number 1, February 2010, 1-16.

[3] Christo Wilson, Bryce Boe, Alessandra Sala, Krishna P.N. Puttaswamy, Ben Y. Zhao, User interactions in social networks and their implications, Proceedings of the 4th ACM European conference on Computer systems, April 01-03, 2009, Nuremberg, Germany.

[4] Ming-Feng Chen, Hsi-Ming Yu, Shu-Fang Li and Ta-Jung You, “A Complementary Method for Detecting Qi Vacuity”, *BMC complementary and alternative medicine*, Vol. 9, No. 12, 2009.

Appendix: Data Analysis and Similarity Retrieval

A slow intelligence system SIS includes many kinds of original data from Tian, Di, Ren and Chi sub-systems. In the TDR system, those data are used to judge a person’s heath status. Two important functions provided by the TRD system are data analysis and similarity retrieval. In data analysis, data will be analyzed so that alert conditions are detected. In similarity retrieval, data records similar to the current data record are found.

1. Data Analysis

The Analyze Model is specified by:(V, V1, V2, V3, VT, VF, VP, VS, VW), where

V is final result, which may be “Normal”, ”Abnormal” or ”N/A”,

V1 is the result of Chi, which may be is “Normal”, ”Abnormal” or ”N/A”,

V2 is the result of bloodpresure, which may be is “Normal”, ”Abnormal” or ”N/A”,

V3 is the result of SPO2, which may be is “Normal”, ”Abnormal” or ”N/A”,

VT is the result of Tongue,

VF is the result of Tongue,

VP is the result of Tongue,

VS is the result of Tongue,

VW is the result of Tongue,

VT, VF, VP, VS, VW have the same value range: 0, 1, 2, 3 or 4.

$$V1 = \begin{cases} Normal & (VT + VF + VP + VS + VW <= 6) \\ Abnormal & (VT + VF + VP + VS + VW > 6) \\ N/A & \left(VT, VF, VP, VS, VW = \text{"}\frac{N}{A}\text{"}\right) \end{cases}$$

$$V2 = \begin{cases} Normal & ((RegularLow - 5 \le Dia \le RegularLow + 5) \text{ and } (RegularHigh - 5 \le Sys \le RegularHigh + 5)) \\ Abnormal & ((RegularLow - 5 \ge Dia \text{ or } Dia \ge RegularLow + 5) \text{ or } (RegularHigh - 5 \ge Sys \text{ or } Sys \ge RegularHigh + 5)) \\ N/A & \left(Dia, Sys = \text{"}\frac{N}{A}\text{"}\right) \end{cases}$$

Where Dia is test result for low pressure, and Sys is for high pressure.

RegularLow and RegulaHigh has a definition as following table.

age	RegularHigh (male)	RegularLow (male)	Regular High (female)	Regular Low (female)
16-20	115	73	110	70
21-25	115	73	110	71
26-30	115	75	112	73
31-35	117	76	114	74
36-40	120	80	116	77
41-45	124	81	122	78
46-50	128	82	128	79
51-55	134	84	134	80
56-60	137	84	139	82
61-65	148	86	145	83

$$V3 = \begin{cases} Normal & (SPO2 \in \{93\% \sim 98\%\}) \\ Abnormal & (SPO2 \text{ not} \in \{93\% \sim 98\%\}) \\ N/A & \left(SPO2 = \text{"}\frac{N}{A}\text{"}\right) \end{cases}$$

V=V1&V2&V3, where ‘&’ is operator, which is defined as follows.

“Normal” & ” Normal”=” Normal”

“Normal” & ”Abnormal”=”Abnormal”

“Normal” & ”N/A”=“Normal”

”Abnormal” & ”Abnormal”=”Abnormal”

”Abnormal” & ”N/A”=”Abnormal”
”N/A” & ”N/A”=”N/A”

The Data Analysis Algorithm is as follows:

（1） Get records from database according to received message,
（2） Get information which includes VT、VF、VP、VS、VW、bloodpresure、SPO2 from all records of the day,
（3） If can’t get enough information at last step, get latest information from the week,
（4） After finished the (3) step, records for some test items still does not exist, those items should be ignore when computing final result.
（5） If there is not any test record in the database, the result is defined as N/A, and insert it into the database, end of algorithm,
（6） According to the formula V=V1&V2&V3, we can get the result(“Normal” or “Abnormal”) and insert into the database, end of algorithm.

2. Similarity Retrieval

This function will retrieve other users’ records, and find out few records that most similar to the current user’s records. This is helpful for user or doctor to make a comparison to enhance his/her understanding of their own situation. For example, if your record is similar to a patient in the last few days, you would better to be aware of your own health condition. In contrast, if your record is similar to a healthy people, it gains your confidence of your health condition.

To find out similar records, the system uses a simplified K Nearest Neighbor algorithm. The system treats each record as a 5 dimensions’ data. The dimensions are Chi factors defined in the system, they are sweaty, pulse, weakBreath, fatigue, and tongue. Each dimension represented by an integer from 1 to 5, that is the score for a specific Chi factor.

After the system retrieve all others records, it calculates the L2 distance (Euclidean distance) between current user’s record and each other’s records using

$$\mathrm{dist}(c, o) = \sqrt{(s_c - s_o)^2 + (p_c - p_o)^2 + (w_c - w_o)^2 + (f_c - f_o)^2 + (t_c - t_o)^2}$$

where c is current user’s record, o is one of other’s record, s is score of sweaty, p is score of pulse, w is score of weakBreath, f is score of fatigue, and t is score of tongue.

When the system knows the distances between current user’s record and others’ records, it can find out the most similar records easily by looking at the distances. Since find similar function do not need to perform any further prediction among current data, thus this can be the end of the nearest neighbor algorithm.

The reason why nearest neighbor algorithm is suitable for this system. One of the reason is, there is almost no assumptions about the data, the only assumptions implied by distance function. In addition, this algorithm is a non-parametric approach, which is mean the data will tell us everything and the system do not need to have any prior knowledge about the data. Although, nearest neighbor algorithm seems fit the system perfectly, it has 2 main disadvantages. One is the algorithm sensitive to irrelevant attributes, also known as dimension curse. Which is mean, high dimension data makes the distance meaningless. Fortunately, the data in the system is 5-dimension data, thus the system do not have this problem. Another disadvantage is the algorithm is computationally expensive, either space wise and time wise. For space wise, the algorithm need to store all examples. However, the system has to keep all the records anyway, therefore, this is not a problem for the system. For time wise, the system has to compute distance to all records with time complexity $O(nd)$, where n is number of existing records, d is cost of computing distance. It is easy to know that, with increasing of n, the system will become slower. The solution in this system is, it only uses recent data so that to control the total number of records involve in the calculation. Due to property of slow intelligence system, data were collected continuously in each cycle, thus missing data will appear in the records. To deal with missing data, the system just simply ignores the incomplete records. To sum up, the nearest neighbor algorithm is one of the most suitable for this system.

Distributing the User Interface Logics along Actionable Components: the EFESTO Approach

Giuseppe Desolda[1], Carmelo Ardito[1], Maristella Matera[2], Maria Francesca Costabile[1]
[1]Dipartimento di Informatica, Università degli Studi di Bari Aldo Moro
Via Orabona, 4 – 70125 – Bari, Italy
{giuseppe.desolda, carmelo.ardito, maria.costabile}@uniba.it

[2]Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano
P.zza L. da Vinci, 32 – 201233 – Milano
maristella.matera@polimi.it

***Abstract* — Developing interactive systems is a very tough task. In particular, the development of user interfaces (UIs) is one of the most time consuming aspects in the software lifecycle. Software development is more and more moving toward composite applications. In this paper, we present a mashup model that enables the integration at the *presentation layer* of specific *UI components*. As application of this model, a mashup platform has been developed that allows non-technical end users to create component-based interactive workspaces via the aggregation and manipulation of data fetched from distributed online resources, also enabling the collaborative creation and use of distributed interactive workspaces. It is shown in this paper how the developed platform permits the rapid prototyping of interactive applications enabling the access to Web services and APIs.**

Keywords - Human-Centric Service Composition; Mashup Model.

I. Introduction

The development of user interfaces (UIs) is one of the most time consuming aspects in the creation of interactive systems. Proper reuse mechanisms for building UIs have become evident, especially as software development is more and more moving toward composite applications [1]. To respond to this need, in this paper we propose a mashup model that enables the integration at the *presentation layer* of "Actionable UI components", which are components equipped with both data visualization templates and a proper logic consisting of functions to manipulate the visualized data. The goal of the model is to reduce the effort required for the development of *interactive workspaces* [2], by maximizing the reuse of UI components.

In our approach, UI components not only constitute "pieces" of UIs that can be assembled together into a unified workspace; each single component can be set to provide views over the huge quantity of data exposed by Web services and APIs available online or by any data source, even personal or locally provided. With respect to the original definition of UI components [3, 4], we promote the notion of Actionable UI component, which also include UI Components that provide varying functions to allow end users to manipulate the contained data.

DOI reference number: 10.18293/DMS2016-034

Our approach can be positioned into the research context related to facilitating the access to data sources through visual user interfaces, a problem that has been attracting the attention of several researchers in the last years [5, 6]. An ever-increasing number of resources that provide content and functions in different formats through programmatic interfaces is available. The efforts of many research projects have thus focused on letting laypeople, i.e., users without expertise in programming, to access and exploit the available content [7, 8, 9]. With this respect, the reuse of easily programmable UI components is a step forwards the provision of environments facilitating the End-User Development (EUD) of service-based interactive workspaces [10]. In general, EUD refers to the involvement of end users in the software development process, in order to modify and even create software artifacts [10, 11, 12]. EUD activities range from simple parameters setting to integration of pre-packaged components, up to extending the system by developing new components. Such a reuse is typical of Web mashups [1], a new class of applications that can be created by integrating components at any of the application stack layers (presentation, business logics, data).

The very novelty introduced by mashups is the possibility to synchronize components equipped with a UI at the presentation layer, for example by means of event-driven composition techniques. Thanks to the possibility of reusing and synchronizing ready-to-use UI components, the mashup has resulted into an effective paradigm to let end users, even non-experts in technology, compose their interactive Web applications. In the last two decades, several mashup tools characterized by different composition paradigms have been proposed. One of the pioneers (even if it is not available anymore) was Yahoo!Pipes [8] that, in the attempt to better support end users, provided a visual editor to access services and operators visualized as visual modules that could be combined into a canvas pane by means of 'pipes' through drag and drop actions. A different paradigm was implemented in NaturalMash, which allows users to express in natural language what service(s) to use and how to synchronize them [7]. However, problems related to the use of a natural language still remain. A different mashup technique avoids the use of APIs and enables data integration allowing users to act directly on the Web page UI elements, which are considered interactive artefacts that can be combined through a set of mashup

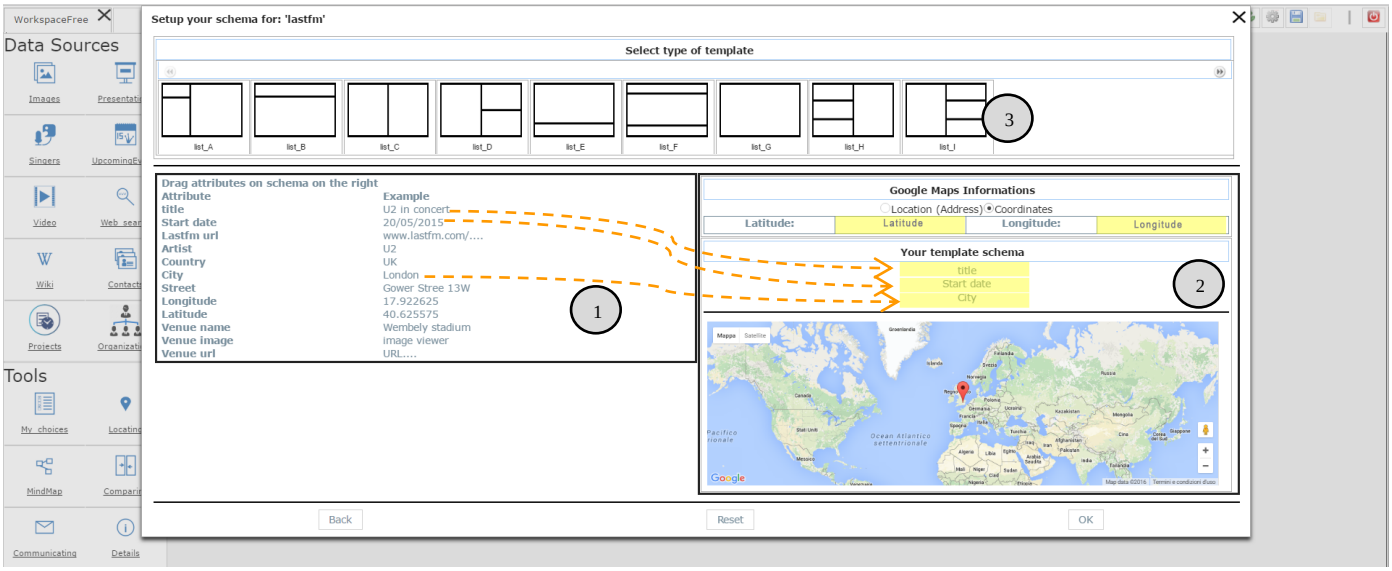


Figure 1. Mapping between some Last.fm attributes and the fields of the *map* user template (circle 2).

operations [13]. The inconvenience of this approach is that it is difficult to generalize due to the heterogeneity of web pages. Other mashup tools and mashup techniques are described in [1]. Moreover, a recent article surveys web service composition techniques and tools [14].

In the last years we have been working extensively on mashup platforms that, by exploiting end-user development principles, address the creation of component-based interactive workspaces by non-technical end users via the aggregation and manipulation of data fetched from distributed online resources [2, 15], also enabling the collaborative creation and use of distributed interactive workspaces [16]. The platform prototype keeps improving on various aspects, based on field studies performed with real users that are revealing new requirements and features useful to foster the adoption of mashup platforms in people daily activities. Based on these experiences, in which we observed people creating easily their interactive applications, in this paper we aim to stress the importance of such platforms as tools for the rapid prototyping of interactive applications enabling the access to Web services and APIs. In particular, the main contribution of this work is a model for UI component mashup that other designers and developers can adopt to develop mashup platforms as tools to easily create interactive workspaces whose logic is distributed across different synchronized components.

The paper is organized as follows: Section II illustrates the main functionality offered by the platform for the creation of interactive workspaces. Section III highlights how the supported modus operandi is made possible thanks to some abstractions, and in particular to the notion of *actionable UI components*, around which the whole platform design has been conceived. In particular, we stress how the adoption of such conceptual elements leads to the notion of *distributed User Interface*, as an interactive artefact that can be assembled according to lightweight technologies and that leverages on the logics of self-contained actionable UI components. Section IV is about Domain-Specific Languages (DSLs) we introduced to describe the main elements of a mashup platform that can guide the dynamic instantiation and execution of the distributed UIs. Section V complements Section III by providing some technical details on how the model elements are implemented in the EFESTO platform architecture. Section VI concludes the paper and outlines future work.

II. The EFESTO Platform

This section describes the most important features of our mashup platform, called EFESTO by showing how it is used to create a mashup. The platform name was inspired by Efesto, a god of the Greek mythology, who realized magnificent magic arms for other Greek gods and heroes. Analogously, the EFESTO platform aims to provide end users with powerful tools to accomplish their tasks. In order to get the reader's attention, such features are written in **bold** in this section and formalized in the model reported in Section III.

A. *Mashup of Data Sources*

In order to describe how EFESTO works, a scenario is reported in which Alan uses the platform to create a mashup that satisfies his information needs. Alan is a non-technical user, i.e., he does not know programming language and he is not familiar with technical terms of computer science.

Alan is going to organize his summer holidays, but he did not yet decide whether to go to Paris or Rome. Regardless the destination, Alan would like to attend at least a concert during his holidays. Thus, he uses EFESTO to create a new application (mashup) that retrieves and integrates information about music events, possibly coming from different sources, and presents the results through a visual representation he selects. Specifically, Alan starts looking for pertinent services among those registered in the platform. A wizard procedure guides him to make a selection from a popup window where services are presented by category (e.g., videos, photos, music, social). Alan clicks on the music category and, among the music services shown, he selects Last.fm, a service that provides information on music events of a specific singer. EFESTO provides different visual templates, called User Interface Templates (**UI Template**), that the user can select in

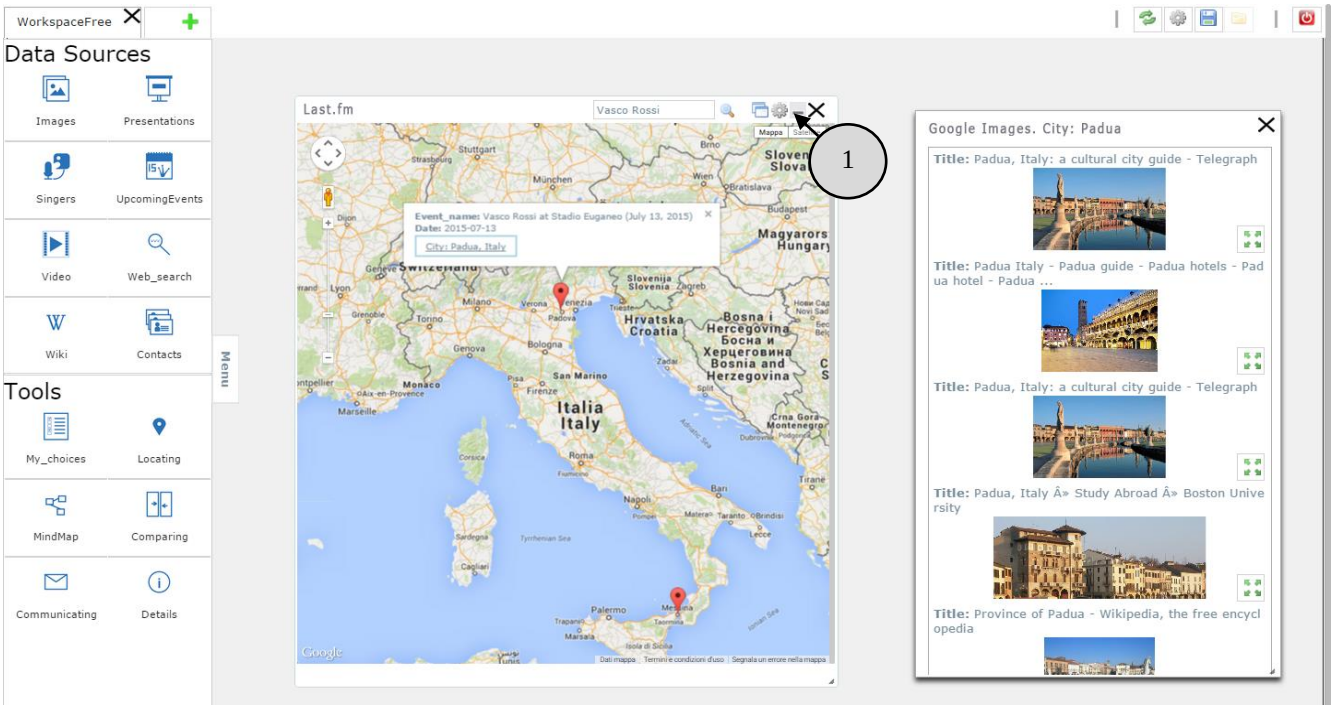


Figure 2. UI Component originated from Last.fm data source visualized as a map and joined with Google Images to show pictures of each city event.

order to display the results of the application he is creating. Alan actually selects a *map* as UI Template, since he wants to visualize the retrieved music events geo-localized in a map.

Among the different data attributes of the Last.fm dataset, Alan has to select those he is interested in, i.e., those that will be considered by the application he is creating. EFESTO enables Alan to make this selection by direct manipulation of elements shown in the user interface of his workspace. In fact, all Last.fm data attributes are visualized in a panel on the left (see Figure 1, circle 1). To make the attributes more understandable, the system also shows some example values. Alan wants his application to consider latitude and longitude of the location where a music event will be performed, so that this location will be visualized in the resulting map. Thus, Alan drags & drops the latitude and longitude Last.fm attributes into the respective fields (called **Visual Renderers** [4]) of the map UI template (Figure 1, circle 2). Alan wants also to visualize, when required, additional details about a musical event. For this, among the available UI templates for text layout (Figure 1, circle 3), he chooses a *table* with three rows and one column, since he wants to visualize three more attributes, namely *title*, *start date* and *city*. To make this possible, he selects each of these three attributes from the left panel (Figure 1, circle 1) and drops it in the visual renderers of the UI template (highlighted in yellow in Figure 1, circle 2).

After performing this mapping phase, Alan saves the mashup in the platform. From now on, this mashup is a **UI Component** in the user workspace, which is immediately executed in the Web browser. This UI Component is called "Last.fm" (by default, it takes the name of the source service but the user can rename it) and is represented as a map, as shown in the central panel in Figure 2. By typing "Vasco Rossi" in the search box (thus making a **query**), the **results set** of forthcoming events of this singer are visualized as pins on the map. The map is shown with a proper zoom level so that all the retrieved events are visualized. This zoom level can obviously be varied by the user. By clicking on a pin representing a music event, details of that event (i.e., the attributes *title*, *start date* and *city*) are shown.

Alan can later update the created mashup by integrating data coming from other **data sources** through **union** and **join** data mashup **operations** [15]. Since a non-technical user is not familiar with union and join operation, EFESTO let the user perform such operations, again, through wizard procedures and drag&drop actions. For example, Alan wants to retrieve more music events than those provided by Last.fm. He then integrates Last.fm with Eventful (another service retrieving music events). Technically, this is a union operation. Alan acts directly on the Last.fm UI component previously created by clicking on the gearwheel icon in the toolbar (pointed by circle 1 in Figure 2) and choosing the "Add results from new source" menu item. A wizard procedure now guides Alan in choosing the new service, Eventful in this example, and in performing a new mapping between the Eventful attributes and the UI template already used in the previous mashup. The newly created mashup (UI Component) is shown in the same fashion reported in Figure 2 but now, when queried with an artist name, this UI Component visualizes results gathered both from the Last.fm and Eventful services.

Another data integration operation available in EFESTO is the join of different sources; it is useful to satisfy user's desire of further integrating the mashup with new data available in other services. For example, Alan would like to show images of the location where music events are held. Last.fm does not provide such images but Alan can retrieve them from Google Images. Technically, this operation is a join between the

Last.fm city attribute and Google Images. EFESTO supports Alan in a very simple way. Alan clicks on the component gearwheel icon and choses the "Extend results with new data" menu item. A new wizard procedure guides him while choosing (a) the service attribute to be extended (City in this example), (b) the new data source (Google Images) and (c) how to visualize the Google Images results. From now on, when clicking on the city name in the map info window, another window visualizes the Google Images pictures related to the selected city, as shown in the right panel of Figure 2.

Another operation available in EFESTO is the *change of visualization* for a given UI component. Alan, in fact, during the interaction with Last.fm, decides to switch from the *map* UI template to the *list* UI template (see the result in Figure 3, circle 1). To perform this action, he clicks on the gearwheel icon in the Last.fm toolbar and chooses the "Change visualization" menu item. A wizard procedure guides Alan to (a) choose a UI template (*list* in this case), and (b) drag&drop the Last.fm attributes onto the UI template, as already described with reference to Figure 1.

B. A polymorphic data source

Despite the wide availability of data sources and composition operations, sometimes users can still encounter difficulties while trying to accommodate different needs and desires. Let us suppose, for example, that during the interaction with EFESTO Alan wants to get details about the artists of the music events, such as genre, starting year of activity and artist photo. Among those services registered in the platform, Alan does not find any that satisfy this new information need. Thus, he should go to the Web for a usual (manual) search for the specific information. However, it might happen that, even on the Web, there are no APIs providing such information.

In order to overcome this drawback, EFESTO provides a new **polymorphic data source** that exploits the wide availability of information structured in the Linked Open Data (LOD) cloud. It is called polymorphic because, when it is composed with another source S, it is capable of modifying its set of attributes depending on the source S, in order to better fulfil the users' needs. In contrast, the standard data sources (YouTube, Wikipedia, etc.) provide the same set of attributes independently of the composing source S. Lack of space prevents us to provide more details about the creation of the polymorphic data source. The interested reader may refer to [17]. DBpedia has been chosen as initial LOD cloud thanks to the vast amount of information it provides.

Thus, Alan can join the Last.fm artist attribute with the DBpedia-based polymorphic data source. The platform now shows a list of attributes related to the musical artist class (available in the DBpedia ontology), and Alan enriches the current UI Component with the attributes *genre*, *starting year of activity* and *artist photo*. Henceforward, Alan can find a list of upcoming events and also visualize artist's information when clicking on the artist's name. What has been described also shows why the DBpedia-based data source is called "polymorphic". In fact, differently from pre-registered data sources (e.g., Google Images) that provide a pre-defined, invariable set of attributes, the system provides users different attributes of the data of the DBpedia-based data source; such attributes are automatically selected depending on the attribute in the origin data source it is bound to. For example, if the Alan's join starting point is the attribute *city*, attributes like *borough*, *census*, *year*, *demographics* would be provided by the DBpedia-based data source.

C. Task-Enabling Containers

Our field studies [2, 16] revealed that mashups generally lack data manipulation functions that end users would like to exploit in order to "act" on the extracted contents, e.g. functions that allow to perform tasks such as collecting&saving favourites, comparing items, plotting data items on a map, inspecting full content details, organizing items in a mind map in order to highlight relationships. In this section, we remark another very innovative feature of EFESTO: it offers tools that enable specific tasks, allowing users to manipulate the information in a novel fashion, i.e., without being constrained to pre-defined operation flows typical of pre-packaged applications.

In order to perform more specific and complex sense-making tasks, a set of Tools are available in the left-panel of the workspace (see Figure 3, circle 4). These Tools are added to the workspace by clicking the corresponding icon. Let us describe an example of their usage with reference to our scenario. Alan is looking for hotels in Rome located nearby the places where upcoming musical events will be held. According to his strategy, he is more interested in finding a good hotel and then look for possible musical events to attend. First, he adds the *Hotel* data source into his workspace (see Figure 3, circle 5) and then performs a search by typing "Rome" in the *Hotel* search bar. After including the *Comparing* tool in the workspace, Alan drags&drops inside it the first five hotels from the *Hotel* UI component. The *Comparing* tool supports Alan in the identification of the most convenient hotels, which are now represented as cards providing further details, such as average price, services and category (see Figure 3, circle 2). Afterwards, he drags&drops three hotels from the *Comparing* tool inside the *Locating* tool (Figure 3, circle 3) in order to visualize them as pins on the map. Finally, Alan performs a search on the *Last.fm* data source by using "Rome" as keyword and then moves all the results, i.e. the upcoming musical events, inside the *Locating* container. The map now shows pins indicating both the hotels and the upcoming musical events in Rome. Alan can now easily identify which musical events are close to the hotels he has previously chosen. However, it could happen that Alan adopts a different strategy. He wants to first identify upcoming musical events and then the hotels nearby. He starts by retrieving musical events with *Last.fm* (see Figure 3, circle 1) and then moves some events inside the *Comparing* tool in order to choose the best ones based on musical genre and artists. Afterwards, he drags&drops some of the compared events inside the *Locating* tool and finally adds into this tool the hotels returned by the *Hotel* data source.

As shown in the previous example, the tools provided in EFESTO allow users to interact with information within dedicated *containers*, which enable specific tasks. Thus, we call them Task-Enabling Containers (TECs). To create such flexible environments, a model has been presented in [18] that

permits easy transition of information between different contexts; this model implements some of the Transformative User eXperience (TUX) principles described in [19, 20].

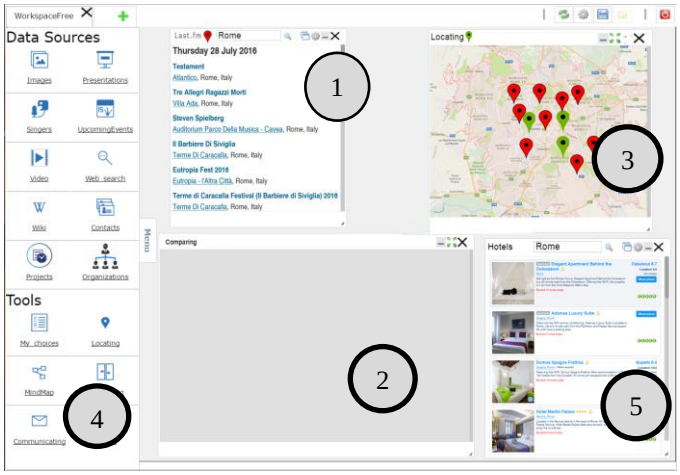


Figure 3. Use of some tools available in EFESTO to manipulate mashup data.

III. Model for UI Component Mashup

The main contribution of this paper is a model highlighting the most important components that make a mashup tool an environment where UIs can be built by reusing and synchronizing the logic of different pieces of UI. The goal is to provide designers and software engineers with a model that guides them during the development of mashup platforms for non-programmers. The proposed model refines and extends the one presented in [4], where the authors defined the modelling abstractions on which their composition paradigm is based. Our model has been iteratively refined by adding further components starting from requirements we gathered

during our research, e.g., a different way to integrate service data by means of more powerful data mashup operations join and union, the integration of some Transformative User eXperience principles [19, 20], and the polymorphic data sources based on Linked Open Data. The new model is depicted in Figure 5. In the following, we report the definitions of the most salient concepts that contribute to the notion of distributed UIs.

Definition 1. *UI Component.* It is the core of the model since it represents the main modularization object the user can exploit to retrieve and compose data extracted from services. It supplies a view according to specific UI Templates (see Definition 2) over one or more services whose data can be composed by means of data mashup operations and also allows the interaction with services data and functions thanks to its own UI (see Figure 4). In addition, two or more UI components can also be synchronized according to an event-driven paradigm: each of them can implement a set E of events that the user can trigger during the interaction with its user interface, and a set A of actions activated when events are performed on others UI components.

Definition 2. *UI Template.* It plays two fundamental roles inside the UI component: first, it guides the users in materializing abstract data sources by means of a mapping between the data source output attributes and the UI template visual renderers; second, at runtime, it displays the data source according to the user mapping. A *UI Template* can be represented as the triple

$$uit = < type, VR >$$

where *type* is the template (e.g., list, map, chart) selected by the user while *VR* is a set of *visual renderers*, i.e., UI elements that act as receptors of data attributes.

Definition 3. *Actionable UI Component (auic).* In addition to visualizing Web service data, *auic* also supply task-related functions for manipulation and transformation of data items retrieved from a source along user-defined task flows [18].

An *auic* can be defined as a pair:

$$auic = <TF, uit>$$

where TF is the set of *functions* for manipulation and transformation of data, while uit is a UI templates used to visualize data according to user's task.

Definition 4. *Event-driven Coupling*. It is a synchronization mechanism among two UI components that the users define according to an event-driven, publish-subscribe integration logic [3]. In particular, the users define that, when an event is triggered on a UI component, an action will be performed by another UI component. This enables reusing the logic of single UI components, still being able to introduce some new behaviour for the composite UIs.

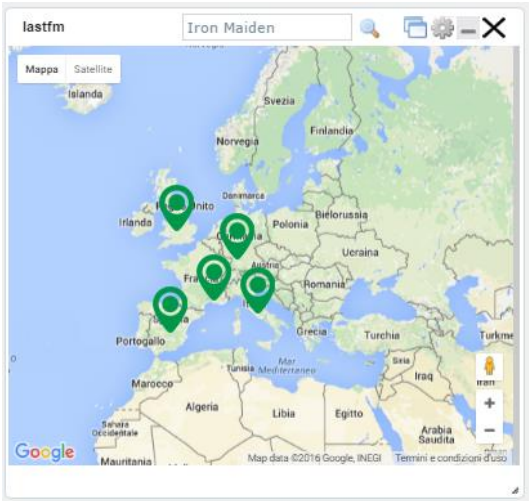


Figure 4: Example of UI Component that shows musical events on Google Maps.

More in general, given two UI Components uic_i and uic_j, a coupling is a pair:

$$c = < uic_i\ (<output>),\ uic_j\ (<input>) >$$

Definition 5. *Presentation Template*. It is an abstract representation of the workspace defining the visual organization of the UICs included in the interactive workspace under construction. For example, the UI components can be freely located or can be constrained to a grid schema, where in each cell only one UI Component can be placed.

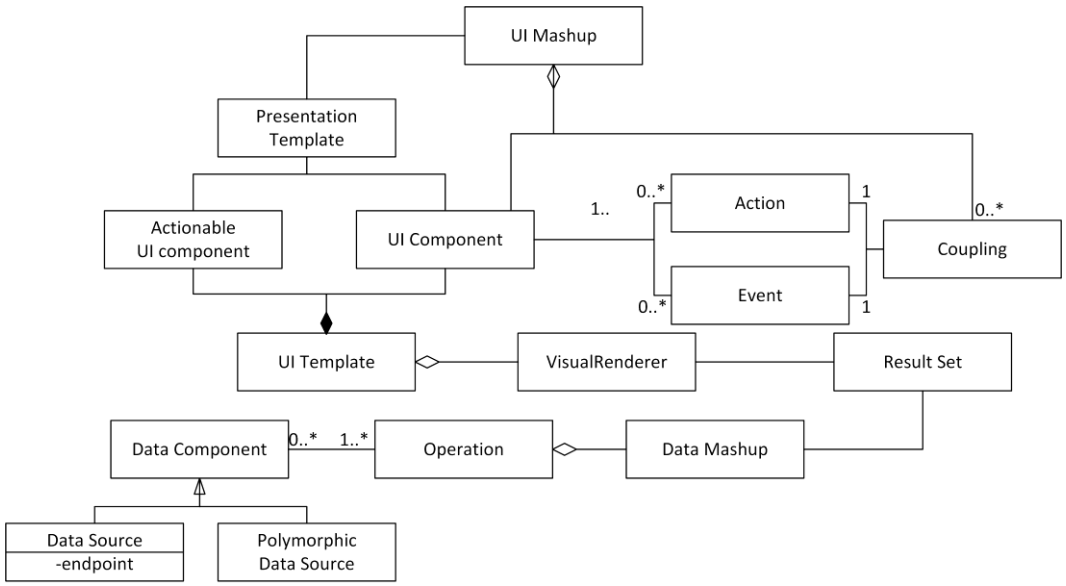


Figure 5: The mashup model.

Definition 6. ***UI Mashup***. A UI Mashup is the final interactive application built by the end users by means of the integration of different UI components within a workspace. It can be formalized as the tuple:

$$UI_Mashup =< UIC, C, PT, AUIC>,$$

where UIC is the set of UI Components integrated into the workspace, C is the set of couplings the users established among UIC, PT is the workspace template chosen to arrange the UIC and AUIC is the set of Actionable UI Components to manipulate data extracted from UIC.

The following definitions are reported to clarify how actionable UI components are instantiated by means of data extracted from data sources.

Definition 7. ***Data Component.*** It is an abstract representation of the resource that can be used to retrieve data. In particular, *dc* is a triplet:

$$dc=<t, I, Out>$$

where *t* indicates the type of resource, for example *REST Data Source* or *Polymorphic Data Source* in our model, *I* indicates the set of input parameters to query the resources, *Out* indicates the set of output attributes. Data can be retrieved from data sources and aggregated through the following operations:

Definition 8.a ***Selection***. Given a data component *dc*, a selection is a unary operator defined as:

$$\sigma_C (dc) = \{r \in dc \mid result\ r\ satisfies\ condition\ C\}.$$

where r is a result obtained by querying the data component dc and C is a condition used to query dc.

Definition 8.b ***Join.*** Given a couple of data components $dc_i =< ep_i, q_i, A>$ and $dc_j =< ep_j, q_j, B>$, a *Join* is a binary operator defined as:

$$dc_i \mid\!><\!\mid_{ai} dc_j = \{(a_1, \ldots, a_n, \sigma_C (dc_j)) \mid C: q_j = a_i \}$$

Definition 8.c ***Union.*** Given a couple of data components $dc_i =< ep_i, q_i, A>$ and $dc_j =< ep_j, q_j, B>$, a *Union* is a binary operator defined as:

$$dc_i \cup dc_j = \{ x \mid x \in dc_i \text{ or } x \in dc_j \}$$

The result of the applying one or more operations is a data mashup, i.e., the composite result set whose rendering and manipulation is possible by means of UI components and Task-Enabling Containers.

Definition 9. ***Data Mashup***. It is the results of the integration of data extracted by different data components. It is a pair:

$$dm=<DC,O>$$

where DC represents the set of data components involved in the composition; O is the set of operations (e.g., join and union) performed between data components in DC.

Data mashup represents an important advance w.r.t. the original model presented in [4] where data mashup was conceived just as a *visual aggregation* of different data sources by means of union and merge sub-templates. In that case, the result of the data mashup could not be reused with other UI templates. In our model, the data mashup is a new integrated result set published as a new data source. This new data source can be used in the platform as a new source that can be visualized by using UI templates.

IV. Platform Descriptors

In order to make the previous abstractions concrete in the implemented platforms, we defined some Domain-Specific Languages (DSLs) inspired to EMML [21]. New languages were adopted instead of EMML because the composition logic implemented in the EFESTO refers only to a small sub-set of the composition operators available in EMML. Each of these

new languages allow us to define internal specifications of the main elements (e.g., UI components, service, UI template) that can guide the dynamic instantiation and execution of the distributed UIs.

In Figure 6**Errore. L'origine riferimento non è stata trovata.** it is reported an example of our XML language specifying a UI component that renders a data mashup consisting in a union between two services (YouTube and Vimeo) and a join of the unified services with a third service (Wikipedia). In the XML file, the tag *unions* has two children, *services* and *shared*. The *services* tag summarizes the unified services. Each service is reported in a *service* tag. In particular, the *service* tag has the attribute *name* that indicates the name of the data source. This value is used by the mashup tool to retrieve the source details to perform the query. The *shared* tag describes the alignment of the attributes of the unified data sources. For example, it has two children called *shared_attribute*, each of them with two children *attribute* that represent the service attributes that are mapped in a *UI template.*

```
<composition join="true" union="true">
 <unions>
  <services>
    <service name="lastfm">
      <attribute name="title" path="title">title</attribute>
      <attribute name="Start date" path="startDate">startDate</attribute>
      <attribute name="City" path="venue.location.city">venue.location.city</attribute>
    <service>
    <service name="songkick">
      <attribute name="EventTitle" path="title">event_title</attribute>
      <attribute name="EventDate" path="startDate">event_date</attribute>
      <attribute name="EventCity" path="venue.location.city">event_city</attribute>
    <service>
  </services>
  <shared>
   <shared_attribute name="title">
    <attribute from_service="lastfm">title</attribute>
    <attribute from_service="songkick">event_title</attribute>
   </shared_attribute>
   <shared_attribute name="Start date">
    <attribute from_service="lastfm">startDate</attribute>
    <attribute from_service="songkick">event_date</attribute>
   </shared_attribute>
   <shared_attribute name="City">
    <attribute from_service="lastfm">venue.location.city</attribute>
    <attribute from_service="songkick">event_city</attribute>
   </shared_attribute>
  </shared>
 </unions>
 <joins>
  <join>
    <service name="youtube" />
    <input>title</input>
    <extendedAttributes>
      <attribute name="Title" path="title">title</attribute>
      <attribute name="Owner" path="path">path</attribute>
    </extendedAttributes>
  </join>
 </joins>
</composition>
```

Figure 6: An example of UI component descriptor codified with our XML language.

Each service reported in the *service* tag is detailed in a separate service descriptor XML file. In Figure 7**Errore. L'origine riferimento non è stata trovata.**, the YouTube service descriptor is reported: inside the root tag called *service*, there are the tags *source*, *inputs*, *params*, *attributes* and *flags.* The first three nodes represent all the information useful to query a data source. The fourth node, *attributes*, describes the instance attributes. The last node, *flag*, is introduced to solve the heterogeneity problem of the data sources. In fact, the remote web services typically send the results by using JSON file but the list of results if formatted in different ways (e.g. inside a JSON array).

```
<service name="youtube" type="JSON" Auth="none">
   <source name="youtube"
           url="https://www.googleapis.com/youtube/v3/search?"
           type="GET">https://www.googleapis.com/youtube/v3/search?
   </source>
   <inputs>
     <input name="q"></input>
   </inputs>
   <params>
     <param name="key">AIzaSyCy5Be6QWdqQ</param>
   </params>
   <separator>&amp;</separator>

   <attributes>
     <attribute name="Title" path="snippet.title">Fear of the Dark</attribute>
     <attribute name="Video" path="id.videoId">Video player</attribute>
     <attribute name="PublicationDate" path="snippet.publishedAt">2009-05-12</attribute>
   </attributes>
   <flags>
     <flag name="is_array">false</flag>
     <flag name="array_name">$.items[*]</flag>
   </flags>
</service>
```

Figure 7: An example of service descriptor codified with our XML language.

Another XML descriptor introduced in our model regards the UI Template. In Figure 8, the *list* UI Template has been reported. It is characterized by a set of sub-UI templates (different types of lists). In particular, the root node, *template*, has an attribute name that indicates the template name. The root has a set of children that describe different alternatives to visualize the UI template.

```
<template type="list">
  <sub_template name="list_A">
    <column width="30%">
      <component name="0" height="20%"/>
      <component name="1" height="80%"/>
    </column>
    <column width="70%">
      <component name="2" height="100%"/>
    </column>
  </sub_template>
  <sub_template name="list_B">
    <column width="100%">
      <component name="0" height="20%"/>
      <component name="1" height="80%"/>
    </column>
  </sub_template>
</template>
```

Figure 8: An example of list UI template descriptor codified with our XML language.

The UI template descriptor is linked with the VI schema through the XML mapping descriptor. An example of mapping is reported in Figure 9. In this descriptor, the root node, *mappings*, has two attributes: *templatetype* and *templatename*. The first one recalls the name of a UI Template (e.g. *list*), the second one the name of its sub-template (*list_A*).

```
<mappings template_type="list" sub_template="List_A">
  <mapping vr="0" attribute="Title" path=".Title">
  <mapping vr="0" attribute="Author" path=".Author">
  <mapping vr="0" attribute="Video" path=".Video">
</mappings>
```

Figure 9: An example of mapping descriptor

V. From the Model to the Platform Architecture

The model presented in Section III guides designers and software engineers in developing mashup platforms targeting to non-programmers. That model highlights the main concepts of a mashup platform without emphasizing technical aspects. In this section, we report a high-level overview of the architecture of the EFESTO mashup platform, in order to illustrate how it implements the mashup model.

The architecture is characterized by tree-layers (Figure 10). On top there is the *UI layer* whose main feature is a visual language that allows end users to perform mashups without requiring technical skills. Such language is based on *UI Components* that use *UI Templates* and *Actionable UI Components* to allow users to visualize and manipulate data extracted from remote sources. *UI layer* runs in the user's Web browser and communicates with the *Logic* and *Data layer* that run on a remote Web server.

The *Logic Layer* implements components that translate the actions performed by end users at the *Interaction Layer* into the mashup executing logic. In particular, the *Mashup Engine* is invoked each time an event, requiring the retrieval of new data or the invocation of service operations, is generated. The *Event Manager*, instead, manages the UI Components coupling. In particular, when users define a synchronization between two UI Components A and B, it instantiates a listener that waits for an event on A that, when triggered, causes the execution of an action on B, according to the coupling defined by the user.

The *Data Layer* stores the XML-based descriptors described in Section IV into proper repositories. In addition, at this layer there are the remote data sources that reside on different Web servers.

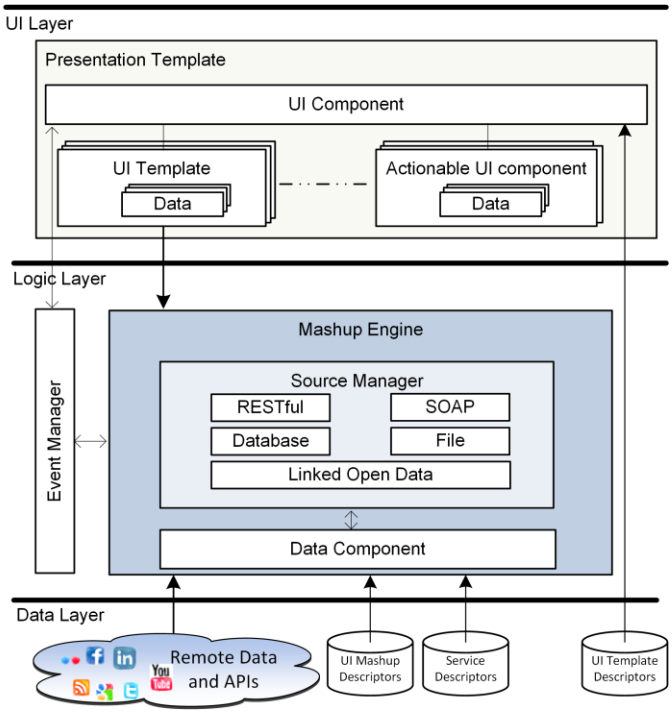


Figure 10. An high-level overview of the EFESTO three-layer architecture

VI. Conclusion

This paper discusses some abstractions that can promote mashup platforms as tools to easily create (i.e., even by end users) interactive workspaces whose logic is distributed across different components that are however synchronized with each other. One of the main contributions of mashup development is the introduction of novel practices enabling integration at the presentation layer in a component-based fashion - an aspect that was never investigated before. Some papers, indeed, discuss and motivate the so-called UI-based integration [4, 22, 23] as a new component-based integration paradigm that privileges the creation of full-fledged artifacts, also equipped with UIs, in addition to the traditional service and data integration practices that instead mainly act at the logic and data layers of the application stack. Along this direction, this paper highlights how interactive artifacts can be composed by reusing the presentation logics (i.e., the UIs) and the execution logics of self-contained modules, the so-called UI Components and Actionable UI Components, providing for the visualization of data extracted by data sources and for data manipulation operations through task-related functions.

The paper capitalizes on the experience gained in the last years by the authors in the development of prototypes of mashup platforms, but it also aims to propose a systematic view on concepts and techniques underlying mashup design and on the way such concepts materialize into composition paradigms and architectures of corresponding development tools, independent of specific approaches and technologies and, thus, of more general validity. Our current work is devoted to enrich the EFESTO platform by means of tools for actionable components and to customize the platform to other application domains, so that further validation studies will be performed.

I. References

[1] Daniel, F., Matera, M.: Mashups: Concepts, Models and Architectures. Springer (2014)

[2] Ardito, C., Costabile, M.F., Desolda, G., Lanzilotti, R., Matera, M., Piccinno, A., Picozzi, M. (2014). User-Driven Visual Composition of Service-Based Interactive Spaces. In *Journal of Visual Languages & Computing* 25(4), 278-296

[3] Yu, J., Benatallah, B., Saint-Paul, R., Casati, F., Daniel, F., Matera, M. (2007). A framework for rapid integration of presentation components. In: *Proc. of WWW '07*. Banff, Alberta (Canada). May 8-12. pp. 923-932

[4] Cappiello, C., Matera, M., Picozzi, M. (2015). A UI-Centric Approach for the End-User Development of Multidevice Mashups. In *ACM Transaction Web* 9(3), 1-40

[5] Spillner, J., Feldmann, M., Braun, I., Springer, T., Schill, A. (2008). Ad-Hoc Usage of Web Services with Dynvoker. In Mähönen, P., Pohl, K., Priol, T. *Towards a Service-Based Internet - ServiceWave 2008*. (Vol. 5377, pp. 208-219)

[6] Krummenacher, R., Norton, B., Simperl, E., Pedrinaci, C. (2009). SOA4All: Enabling Web-scale Service Economies. In: *Proc. of ICSC '09*. Berkeley, CA (USA). 14-16 September. pp. 535-542

[7] Aghaee, S., Pautasso, C. (2014). End-User Development of Mashups with NaturalMash. In *Journal of Visual Languages & Computing* 25(4), 414-432

[8] Pruett, M.: Yahoo! pipes. O'Reilly (2007)

[9] Hirmer, P., Mitschang, B. (2016). FlexMash–Flexible Data Mashups Based on Pattern-Based Model Transformation. In Daniel, F., Pautasso, C. *Rapid Mashup Development Tools - Rapid Mashup Challenge in ICWE 2015*. (Vol. 591, pp. 12-30)

[10] Fischer, G. (2009). End-User Development and Meta-design: Foundations for Cultures of Participation. In Pipek, V., Rosson, M.B., de

Ruyter, B., Wulf, V. *End-User Development - Is-EUD 2009*. (Vol. pp. 3-14)

[11] Costabile, M.F., Fogli, D., Mussio, P., Piccinno, A. (2006). End-User Development: The Software Shaping Workshop Approach. In Lieberman, H., Paternò, F., Wulf, V. *End User Development - Is-EUD 2006*. (Vol. 9, pp. 183-205)

[12] Costabile, M.F., Dittrich, Y., Fischer, G., Piccinno, A.: End-User Development. Vol. LNCS 6654, Springer-Verlag, Berlin / Heidelberg, 2011 (2011)

[13] Daniel, F. (2015). Live, Personal Data Integration Through UI-Oriented Computing. In: *Proc. of ICWE '15*. Rotterdam, the Netherlands. 23 – 26 June 2015. pp. 479-497

[14] Lemos, A.L., Daniel, F., Benatallah, B. (2015). Web Service Composition: A Survey of Techniques and Tools. In *ACM Computing Survey* 48(3)**,** 1-41

[15] Desolda, G., Ardito, C., Matera, M. (2015). EFESTO: A platform for the End-User Development of Interactive Workspaces for Data Exploration. In Daniel, F., Pautasso, C. *Rapid Mashup Development Tools - Rapid Mashup Challenge in ICWE 2015*. (Vol. 591, pp. 63 - 81)

[16] Ardito, C., Bottoni, P., Costabile, M.F., Desolda, G., Matera, M., Picozzi, M. (2014). Creation and Use of Service-based Distributed Interactive Workspaces. In *Journal of Visual Languages & Computing* 25(6)**,** 717-726

[17] Desolda, G. (2015). Enhancing Workspace Composition by Exploiting Linked Open Data as a Polymorphic Data Source. In Damiani, E., Howlett, J.R., Jain, C.L., Gallo, L., De Pietro, G. *Intelligent Interactive Multimedia Systems and Services (KES-IIMSS '15)*. (Vol. pp. 97-108)

[18] Ardito, C., Costabile, M.F., Desolda, G., Latzina, M., Matera, M. (2015). Making Mashups Actionable Through Elastic Design Principles. In Díaz, P., Pipek, V., Ardito, C., Jensen, C., Aedo, I., Boden, A. *End-User Development - Is-EUD 2015*. (Vol. LCNS 9083, pp. 236-241)

[19] Latzina, M., Beringer, J. (2012). Transformative user experience: beyond packaged design. In *interactions* 19(2)**,** 30-33

[20] Beringer, J., Latzina, M. (2015). Elastic workplace design. In *Designing Socially Embedded Technologies in the Real-World*. (Vol. pp. 19-33)

[21] Viswanathan, A. (2010). Mashups and the enterprise mashup markup language (EMML). In *Dr. Dobbs Journal*

[22] Daniel, F., Matera, M., Yu, J., Benatallah, B., Saint-Paul, R., Casati, F. (2007). Understanding ui integration: A survey of problems, technologies, and opportunities. In *Internet Computing, IEEE* 11(3)**,** 59-66

[23] Yu, J., Benatallah, B., Casati, F., Daniel, F. (2008). Understanding Mashup Development. In *IEEE Internet Computing* 12(5)**,** 44-52

Efficient Unsupervised Behavioral Segmentation of Human Motion Capture Data

Xiaomin Yu[1], Weibin Liu*[1] and Weiwei Xing[2]

[1]Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China
[2]School of Software Engineering, Beijing Jiaotong University, Beijing 100044, China
E-mail:wbliu@bjtu.edu.cn

Abstract *-With the development of human motion capture, realistic human motion capture data has been widely implemented to many fields. However, segmenting motion capture data sequences manually into distinct behavior is time-consuming and laborious. In this paper, we introduce an efficient unsupervised method based on graph partition for automatically segmenting motion capture data. For the N-Frame motion capture data sequence, we construct an undirected, weighted graph G=G(V,E), where the node set V represent frames of motion sequence and the weight of the edge set E describes similarity between frames. In this way, behavioral segmentation problem on motion capture data may be transformed into graph cut problem. However, the traditional graph cut problem is NP hard. By analyzing the relationship between graph cut and spectral clustering, we apply spectral clustering to the NP hard problem of graph cut. In this paper, two methods of spectral clustering, t-nearest neighbors and the Nystrom method, are employed to cluster motion capture data for getting behavioral segmentation. In addition, we define an energy function to refine the results of behavioral segmentation. Extensive experiments are conducted on the dataset of multi-behavior motion capture data from CMU database. The experimental results prove that our novel method is robust and effective.*

Keywords: *motion capture data, behavioral segmentation, graph cut, spectral clustering*

I. INTRODUCTION

Human motion capture, once associated with producing special effects for film and television production, entertainment games and virtual reality, is common today in diverse applications ranging from health care to consumer electronics. The data of human motion capture can be extensively applied in many fields, such as producing realistic animation movies [1][2],physical rehablitation [3][4]and ergonomic analysis[5]. However, the cost of capturing motion data is very high. It's necessary to better reuse motion capture data. Motion capture data sequence is usually comprised of multiple types of behaviors. The present work suggests that automatic segmentation of human motion capture data into distinct behavior based on statistical properties of the motion can be an efficient and quite robust alternative to hand segmentation. This paper focuses on efficient and robust technique which can be able to automatically segment motion capture data sequence into distinct behaviors, as depicted in Figure 1. The main contributions of this paper can be concluded as follows:

(1) Constructing N-Frame motion capture data sequence into an undirected, weighted graph G=G(V,E), which is based on the similarity of data frames. According to this way, we transform motion capture data segmentation problem into the problem of graph cut.

(2) We employ an algorithm, which can automatically extract the number of behavior and cluster centers, to the motion capture data. It's based on an unsupervised cluster method. We consider the cluster numbers as the behavior numbers for motion capture data sequence.

(3) Analyzing in detail for graph cut method and spectral clustering, we transform this NP hard problem into eigenvalues and eigenvectors in spectral space. Two methods of spectral clustering, t-nearest neighbors and the Nystrom method,are employed to cluster motion capture data for getting behavioral segmentation.

(4) For the clustering fragments, we define energy function and use dynamic programming to refine the results of behavioral segmentation for motion capture data.

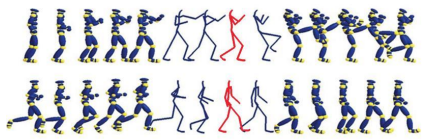


Figure 1: Behavior segmentation. Segmenting motion capture data into distinct behavior.

*Corresponding author: Weibin Liu, wbliu@bjtu.edu.cn

DOI reference number: 10.18293/DMS2016-016

The remainder of this paper is organized as follows. Firstly, we introduce works of predecessors for motion segmentation in Section II. In Section III, how to calculate similarity of motion capture data frames is adopted. The details for identifing behavior's clusters centers for original motion capture data sequences are presented in Section IV. In Section V, we analyze in detail for the relationship between graph cut and spectral clustering. Then, transforming this NP hard problem of graph cut into spectral clustering problem. In this way, realizing behavioral segmentation for motion capture data sequence. Details of energy function for refining the results of behavioral segmentation is represented in Section VI. In section VII, we introduce the experiments, which are conducted on the dataset of multi-behavior motion capture data from CMU database. Finally, we provide the discussion and conclusion for this behavior segmentation method in Section VIII.

II. RELATED WORK

Temporal segmentation is related to numbers of different fields such as data mining [6][7], behavior recognition [8] and so on. Researchers have proposed several techniques to segment motion capture data into distinct behaviors. Maybe they focus on different perspectives and concerns, their purpose is to establish efficient and robust segmentation methods for motioncapture data. The methods can be generally categorized into several types: classifier, clustering, machine learning and dimension reduction and so on.

Supervised learning methods generally formulize segmentation motion capture data as a classification problem, where classifiers are trained from a carefully selected training set. In other words, Its mean that they are usually applying example segments or pre-computed templates and matching them to test sequences. Muller et al.[9]constructed motion templates to behaviors segmentation. Lv et al. [10]defined Hidden Markov Models (HMM) to realize human Motion capture data segmentation. Support Vector Machine(SVM) which is based on an annotation training database to segment motion capture data was constructed by Arikan et al. [11]. However, these methods relied on the training set and they would fail to pick up the segments whose corresponding behaviors were not contained in the training set.

It's natural for researchers to use another technique to overcome this limitation, which can be named as unsupervised learning methods. In these methods, motion capture data segmentation is located by clustering motion frames. zhou et al.[12]used aligned cluster analysis (ACA) to temporally cluster poses into motion primitives which were then assigned to different behavior classes. ACA extends standard kernel k-means clustering: the cluster means include a number of features and a dynamic time warping (DTW) kernel is used to achieve temporal invariance. However, ACA method needs users determine the cluster number with respect to temporal constraint. In order to overcome this limitation, zhou et al. [13] derived an unsupervised hierarchical bottom-up framework, which is called hierarchical aligned cluster analysis (HACA) to realize segmentation. HACA provided a crude method to find a low-dimensional embedding for the time series. HACA is efficiently optimized with a coordinate descent strategy and dynamic programming.

The researchers tried to solve the dilemma of non-learning way to behavior segmentation. Balazia et.al [14] introduced an unsupervised key-pose detection algorithm for segmentation of motion capture data and this proposed algorithm partition motions at the level of gestures. Barbic et al. [15] chose segments using an indication of intrinsic dimensionality from Principal Component Analysis (PCA). Its based on the observation that simple motions exhibit lower dimensionality than more complex motions. As an extension of the traditional PCA, Barbic et al. defined a proper probability model for PCA which is named probabilistic PCA (PPCA). We can easily know that the directions outside the subspace were discarded, whereas they were modeled with noise in PPCA.

III. SIMILARITY MEASURE OF MOTION CAPTURE DATA

A. Distance of Frames

This paper employs the human skeleton model which has 31 joints, as illustrated in Figure 2. For every frame, it has 62-dimensions, which includes root position vector, root orientation vector and other joints' direction vector. The ith frame's pose consists of all joints rotation angle in the ith frame expect the root position vector and the root orientation vector which including 6-dimensional. Each pose $p_i = \{a_{i,1}, a_{i,2}, a_{i,3} \ldots a_{i,56}\}$ is represented as a point in 56-dimensional,which $a_{i,j}$ is one of an Euler angle. The velocity v_i of the ith frame is computed by the Euclidean distance between p_i and p_{i+1}.

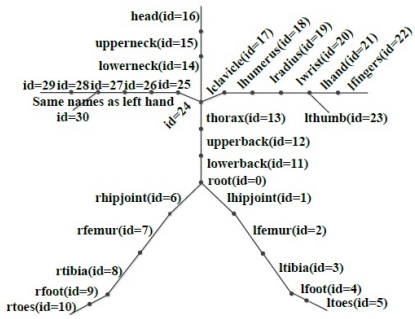


Figure 2: Human Skeleton Model

$$v_i = \begin{cases} \sqrt{(a_{i+1,1} - a_{i,1})^2 + \ldots + (a_{i+1,56} - a_{i,56})^2} & i \neq n \\ v_{i-1} & i = n \end{cases} \quad (1)$$

Calculating the distance by:

$$d_{i,j} = \alpha d(p_i, p_j) + \beta d(v_i, v_j) \quad (2)$$

Where $d(v_i, v_j)$ is the difference of velocity between the ith frame and the jth frame and $d(p_i, p_j)$ is the weighted difference of joint orientations. The weights of α and β are set to 0.5. The term of $d(p_i, p_j)$ is given by:

$$d(p_i, p_j) = ||p_{i,0} - p_{j,0}||^2 + \sum_{k=1}^{m} w_k ||log(q_{j,k}^{-1} q_{i,k})||^2 \quad (3)$$

In (3), $p_{i,0}, p_{j,0} \in R^3$ are the global translational positions of the figure at frame i and j, respectively; m is the number of joints; and $q_{i,k}, q_{j,k} \in S^4$ are the orientations of joint k and frames i and j, respectively. The log-norm term represents the geodesic norm in quaternion space. w_k describes the weight of k-th joint. Lee et al [16] set the weights manually. While the weights of unimportant joints are set to zero, the weights of important joints, such as shoulders, elbows, hips, knees, hips and so on, are set to one. For different joint, it's well known to occupy different role. In this way, there will be a larger error. Wang et al [17] compute a set of optimal weights for the cost function using a constrained least-squares technique, where the weights are calculated in two ways: through a cross-validation study and a medium-scale user study. In order to reduce the error, we use the optimal weights proposed by Wang and Bodenheimer, which are shown in table 1.The weights of remaining joints are set to zero. According to this, we can get a distance matrix D_{n*n} , where n is the length of the original motion capture data sequence. Apparent that $d_{ij} = d_{ji}(i \neq j)$ and $d_{ij} = 0(i = j)$.

B. Frame Kernel Matrices

In general, there are several methods which are used to indicate the similarity of frames for motion capture data. There are several different formulas to define it. In order to better construct the similarity matrix, we optimized the novel representation method of frames similarity. Calculating the similarity of frames by:

$$w_{ij} = e^{(-\frac{dist_{ij}^2}{2\sigma_i \sigma_j})} \quad (4)$$

It's time-consuming to get an optimum parameter σ_i for this function by multiple experiments. In order to determine the value of parameter σ_i, a parameter construction method of neighborhood adaptive scale [18] is introduced:

Table 1: Joints with non-zero weights.

Joints	Weight
Right and Left Hip	1.0000
Right and Left Knee	0.0901
Right and Left Shoulder	0.7884
Right and Left Elbow	0.0247

$$\sigma_i = \frac{1}{k} \sum_{m=1}^{K} D_{ii_m} = \frac{1}{k} \sum_{m=1}^{K} ||i - i_m|| \quad (5)$$

Setting the k that the average number of neighbors is around 1% of the total frame number for every motion capture data sequences. The distance matrix between frames is computed by (2).

IV. IDENTIFY CENTERS OF BEHAVIORAL CLUSTERS

In order to automatically extract the number of motion categories for motion capture data sequence, we use this algorithm, which has good effect for any data points shape. As always, this algorithm has its basis only in the distance between data points. It's able to detect non-spherical clusters and automatically find the correct number of clusters.

This algorithm has its basis in the assumption that cluster centers are surrounded by neighbors with lower local density and that they are at a relatively large distance from any point with a higher local density. It's clear that the same motion behavior with sufficient similarity. We consider every motion capture data frame as a high-dimensional point. For each high-dimensional point, we should compute two quantities: its local density ρ_i and its distance δ_i from points of higher density.

Rodriguez et al. [19] introduce two methods for computing local density:

$$\rho_i = \sum_{j \in I_s \setminus \{i\}} X(d_{ij} - d_c) \quad (6)$$

$$X(x) = \begin{cases} 1 & x < 0 \\ 0 & x \geq 0 \end{cases} \quad (7)$$

d_c is a cut-off distance. ρ_i describes the number of points, which the distance with i is smaller than d_c . Obviously, ρ_i is discrete.

We can see that formula ρ_i has a significant limitation of different motion capture data points with the same local density. In order to avoid this limitation, we calculate ρ_i by:

$$\rho_i = \sum_{j \in I_s \setminus \{i\}} e^{-(\frac{d_{ij}}{d_c})^2} \quad (8)$$

Which is satisfied with more data points of the distance with i smaller than d_c and the bigger ρ_i . δ_i is measured by computing the minimum distance between point i and any other points with higher density:

$$\delta_i = \min_{j:\rho_j > \rho_i}(d_{ij}) \quad (9)$$

The motion capture data frames which are satisfied the conditions of cluster centers should have the larger ρ and δ. Computing $\alpha = \rho * \delta$ and choosing the larger α , which are more likely to become cluster centers. Because of the magnitude of ρ and δ is uniform, it needs to be normalized.

In order to extract cluster centers, we observe the discrete derivative $\Delta_i = |\alpha_i - \alpha_{i-\theta}|$, where θ must be enough to avoid noise in the data. For$\{\alpha_i\}_{i=1}^{N}$, we calculate the average $\overline{\Delta}_i$ and standard derivation ε_i of all data Δ_i . So we extract the point as cluster center when $\Delta_i > 3\varepsilon_i$. we automatically extract cluster numbers k and cluster centers $\{c_1, c_2, c_3, \ldots, c_k\}$ for motion capture data sequences. In this way, we extract k as the number of behavior for motion capture data sequence.

V. BEHAVIORAL SEGMENTATION BASED ON SPECTRAL CLUSTERING

Yuan et al.[20] proposed a graph partition model with temporal constraints to perform temporal segmentation problem. As we all know, the problem of motion capture data segmentation is typical temporal data segmentation. So we base on graph partition theory to segment motion capture data. Giving a N-Frame motion capture data sequence $\{f_1, f_2, f_3, \ldots, f_N\}$, we construct an undirected, weighted graph G=G(V,E). Each node in the set V describes a frame and frames is connected by edges. The similarity between frames represents the weight of each edge. This problem for Segmenting motion capture data can be restated as follows: we want to find a partition of the graph such that the edges between different sub-graphs have a very low weight (which means that points in different clusters are dissimilar from each other) and the edges within the same group have high weight (which means that points within the same cluster are similar to each other).

A. Building Graph for Motion Capture Data

For the N-frame motion capture data sequence, we construct G(V,E) as an undirected graph with vertex set $V = \{f_1, f_2, f_3, \ldots, f_N\}$. In the following we assume that the graph G is weighted. Each edge between two vertices f_i and f_j carries a non-negative weight $w_{ij} \geq 0$. The matrix $W = (w_{ij})_{i,j=1,2,3,\ldots,N}$ represents the weighted adjacency matrix of the graph. When $w_{ij} = 0$ means that the vertices f_i and f_j are not connected by an edge. Because of G is an undirected graph, we define $w_{ij} = w_{ji}$. The degree of a vertex $f_i \in V$ is defined as:

$$d_i = \sum_{j=1}^{N} w_{ij} \tag{10}$$

We construct the degree matrix D, which is defined as the diagonal matrix with the degrees $d_1, d_2, d_3, \ldots, d_N$ on the diagonal. Defining an indicator vector $I = (I_1, I_2, I_3, \ldots, I_N)^{'} \in R^N$ as the vector with entries $I_i = 1$ if $I_i \in F$ (F means subset of vertices $F \subset V$) and $I_i = 0$ otherwise.

For two disjoint sets $F_1, F_2 \subset V$, we define:

$$W(F_1, F_2) = \sum_{i \in F_1, j \in F_2} w_{ij} \tag{11}$$

Considering two different ways of measuring the size of the subset $F_c \subset V$.

$< 1 >$ $|F_c|$ = the number of vertices in F_c;

$< 2 >$ $vol(F_c) = \sum_{i \in F_c} d_i$.

Intuitively, $|F_c|$ measures the size of F_c by its number of vertices, while $vol(F_c)$ measures the size of F_c by summing over the weights of all edges attached to vertices in F_c. The nonempty sets $F_1, F_2, F_3, \ldots, F_k$ form a partition of the graph if $F_{ci} \cap F_{cj} = \emptyset$ and $F_1 \cup \ldots \cup F_k = V$.

B. Construt Graph Laplacians

For the undirected graph G(V,E), we define its Laplacian matrix:

$$L = D - W \tag{12}$$

The Laplacian matrix satisfies the following properties:

$< 1 >$ L has m non-negative, real-valued eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_m$;

$< 2 >$ For every vector $I \in R^N$, we have $I^{'}LI = \frac{1}{2} \sum_{i,j=1}^{N} w_{ij}(I_i - I_j)^2$.

The proof of it as follows:

$$\begin{aligned} I^{'}LI = I^{'}(D-W)I &= \sum_{i=1}^{N} d_i I_i^2 - \sum_{i,j=1}^{N} I_i I_j w_{ij} \\ &= \frac{1}{2}(\sum_{i=1}^{N} d_i I_i^2 - 2\sum_{i,j=1}^{N} I_i I_j w_{ij} + \sum_{j=1}^{N} d_j I_j^2) \\ &= \frac{1}{2}\sum_{i,j=1}^{N} w_{ij}(I_i - I_j)^2 \end{aligned} \tag{13}$$

For N-Frame motion capture data sequence $\{f_1, f_2, \ldots, f_N\}$, we consider every frame as data point for the undirected graph. It is generally known that several popular constructions to transform data points with pairwise similarities w_{ij} or pairwise distances d_{ij} into a graph. When constructing similarity graphs the goal is to model the local neighborhood relationships between data points.

Considering two different ways of constructing similarity graph:

$< 1 >$ t-nearest neighbor graphs: The goal is to connect vertex F_i with vertex F_j if F_j is among the t-nearest neighbors of F_j . We set the t that the average number of neighbors is around 2% of the total number for every motion capture data sequences.

$< 2 >$ The fully connected graph: we use the Gaussian similarity function, which is defined in (4). The parameter σ controls the width of the neighborhoods.

C. Apply Spectral Clustering to Graph Cut

a. Relationship Between Graph Cut and Spectral Clustering

For this undirected and weighted graph, the mincut approach simply consists of choosing a partition $F_1, F_2, F_3, \ldots, F_k$ which minimizes:

$$cut(F_1, F_2, F_3, \ldots, F_k) = \frac{1}{2}\sum_{i=1}^{k} W(F_i, \bar{F}_i) \tag{14}$$

Where $\bar{F}$ for the complement of F. Introducing the factor 1/2 for notational consistency, otherwise we would count each edge twice in the cut. In order to make the weight of cuts edges, which are minimum, it's to make the above objective function minimum.

There are two most common objective functions to encode this are RatioCut, which is defined by Hagen et al.[21], and the nor-

malized cut (Ncut), which is proposed by Shi et al[22].

$$\begin{aligned} RatioCut(F_1,\ldots,F_k) &= \frac{1}{2}\sum_{i=1}^{k}\frac{W(F_i,\bar{F}_i)}{|F_i|} \\ &= \sum_{i=1}^{k}\frac{cut(F_i,\bar{F}_i)}{|F_i|} \end{aligned} \tag{15}$$

$$\begin{aligned} Ncut(F_1,\ldots,F_k) &= \frac{1}{2}\sum_{i=1}^{k}\frac{W(F_i,\bar{F}_i)}{vol(F_i)} \\ &= \sum_{i=1}^{k}\frac{cut(F_i,\bar{F}_i)}{vol(F_i)} \end{aligned} \tag{16}$$

Choosing a motion capture data sequence,which has two behavior, and we analyze the relationship between graph cut and spectral clustering.

The goal of us is to solve the optimization problem:

$$\min_{F\subset V} RatioCut(F,\bar{F}) \tag{17}$$

Defining the vector $I=(I_1,I_2,I_3,\ldots,I_N)^{'}\in R^N$ and getting

$$I_i=\begin{cases}\sqrt{|\bar{F}|/|F|} & if \quad f_i\in F\\ -\sqrt{|F|/|\bar{F}|} & if \quad f_i\in \bar{F}\end{cases} \tag{18}$$

According to the property of Laplacian matrix, which is $I^{'}LI=\frac{1}{2}\sum_{i,j=1}^{N}w_{ij}(I_i-I_j)^2$, we can get:

$$\begin{aligned} I^{'}LI &= \frac{1}{2}\sum_{i,j=1}^{N}w_{ij}(I_i-I_j)^2 \\ &= \frac{1}{2}\Big(\sum_{i\in F,j\in\bar{F}}w_{ij}\Big(\sqrt{\frac{|\bar{F}|}{|F|}}+\sqrt{\frac{|F|}{|\bar{F}|}}\Big)^2 \\ &+ \sum_{i\in\bar{F},j\in F}w_{ij}\Big(-\sqrt{\frac{|\bar{F}|}{|F|}}-\sqrt{\frac{|F|}{|\bar{F}|}}\Big)^2\Big) \\ &= cut(F,\bar{F})\big(\frac{|F|+|\bar{F}|}{|F|}+\frac{|F|+|\bar{F}|}{|\bar{F}|}\big) \\ &= |V|*RatioCut(F,\bar{F}) \end{aligned} \tag{19}$$

In other words, we can get that the optimization problem and the Laplacian matrix has a big relationship. It's easy to extend to k subgraphs.

Because of these constraint conditions:

$$\begin{aligned} \sum_{i=1}^{N}I_i &= \sum_{i\in F}\sqrt{\frac{|\bar{F}|}{|F|}}-\sum_{i\in\bar{F}}\sqrt{\frac{|F|}{|\bar{F}|}} \\ &= |F|*\sqrt{\frac{|\bar{F}|}{|F|}}-|\bar{F}|*\sqrt{\frac{|F|}{|\bar{F}|}}=0 \end{aligned} \tag{20}$$

$$I^{'}*\mathbf{1}=\sum_{i=1}^{N}I_i=0 \tag{21}$$

$$||I||^2=\sum_{i=1}^{N}I_i^2=|F|*\sqrt{\frac{|\bar{F}|}{|F|}}+|\bar{F}|*\sqrt{\frac{|F|}{|\bar{F}|}}=N \tag{22}$$

We can get the new optimization problem:

$$\begin{aligned} &\min_{I\in R^N} I^{'}LI \\ &subject \quad to \quad I^{'}*\mathbf{1}=0, \quad ||I||=\sqrt{N} \end{aligned} \tag{23}$$

Assuming that $LI=\lambda I$, at the moment, λ is eigenvalues and I is $L^{'}s$ eigenvectors. Multiplying from the left and the right by $I^{'}$ for $LI=\lambda I$, we can get $I^{'}LI=\lambda I^{'}I$ (where $I^{'}I=N$). Because of N is constant, minimizing the formula $I^{'}LI$ can be replaced by minimizing λ. For the Normalized cut, we define the cluster indicator vector $I=(I_1,I_2,I_3,\ldots,I_N)^{'}\in R^N$ by:

$$I_i=\begin{cases}\sqrt{vol(\bar{F})/vol(F)} & if \quad f_i\in F\\ -\sqrt{vol(F)/vol(\bar{F})} & if \quad f_i\in \bar{F}\end{cases} \tag{24}$$

Likely the RatioCut, we can get:

$$\begin{aligned} I^{'}LI &= \frac{1}{2}\sum_{i,j=1}^{N}w_{ij}(I_i-I_j)^2 \\ &= \frac{1}{2}\Big(\sum_{i\in F,j\in\bar{F}}w_{ij}\Big(\sqrt{\frac{vol(\bar{F})}{vol(F)}}+\sqrt{\frac{vol(F)}{vol(\bar{F})}}\Big)^2 \\ &+ \sum_{i\in\bar{F},j\in F}w_{ij}\Big(-\sqrt{\frac{vol(\bar{F})}{vol(F)}}-\sqrt{\frac{vol(F)}{vol(\bar{F})}}\Big)^2\Big) \\ &= cut(F,\bar{F})\big(\frac{vol(F)+vol(\bar{F})}{vol(F)}+\frac{vol(F)+vol(\bar{F})}{vol(\bar{F})}\big) \\ &= vol(V)*NCut(F,\bar{F}) \end{aligned} \tag{25}$$

$vol(V)$ is a constant. The goal of us is to solve the optimization problem:

$$\min_{F\subset V} NCut(F,\bar{F}) \tag{26}$$

It's equal to $\min_{I\in R^N} I^{'}LI$. In order to deal with this optimization problem, we calculate the Laplacian matrixs minimum eigenvalue. However, the smallest eigenvalue of L is 0 with eigenvector 1. The Rayleigh-Ritz theorem can be used to solve this problem. The solution of this problem is given by the vector I which is the eigenvector corresponding to the second smallest eigenvalue of the Laplacian matrix L . So we can approximate a minimizer of RatioCut or Ncut by the second eigenvector of L . Extending to k clusters, we compute the eigenvalues of the Laplacian matrix and sort them according to order from small to big. Eigenvectors corresponding to eigenvalues are also sort in increasing. Extracting k eigenvectors as what we want. In this way, we succeeded in converting the graph cut, which is a NP problem, into the Laplacian matrix eigenvalues (eigenvectors) problem.

b. Apply Nystrom Method to Spectral Clustering

The $Nystr\ddot{o}m$ method[23], which uses a sub-matrix of the dense similarity matrix: This method is a technique for finding an

approximate eigendecomposition. Here, we denote by W , which is a N*N similarity matrix. Assume that we randomly select sample $l << N$ points from the data. The matrix A represent the $l \times l$ matrix of similarities between the same points, B be the $l \times (n-l)$ matrix of affinities between the l sample points and the $(n-l)$ remaining points, C contains the similarities between all $(n-l)$ remaining points and O be the $(n \times l)$ matrix consisting of A and B^T. We can get rearrange the similarity matrix W such that:

$$W = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \quad and \quad O = \begin{bmatrix} A \\ B^T \end{bmatrix} \tag{27}$$

For the $Nyströ m$ method, it use matrix A and B to approximate similarity matrix W.

$$\begin{aligned} W &\approx \tilde{W} = OA^{-1}O^T \\ &= \begin{bmatrix} A & B \\ B^T & B^T A^{-1} B \end{bmatrix} = \begin{bmatrix} A \\ B^T \end{bmatrix} A^{-1} \begin{bmatrix} A & B \end{bmatrix} \end{aligned} \tag{28}$$

Where the matrix C is now replaced by $B^T A^{-1} B$. The matrix A makes eigendecomposition and we can get $A = U_A \Sigma_A U_A^T$, where Σ_A contains the eigenvalues of A and U_A are the corresponding eigenvectors. According to the $Nyströ m$ method, we can get :

$$\tilde{\Sigma} = (\frac{N}{l})\Sigma_A, \tilde{U} = \sqrt{\frac{l}{N}} O U_A \Sigma_A^{-1} \tag{29}$$

Moreover, the similarity matrix W has the eigendecomposition:

$$\tilde{W} = \tilde{U}\tilde{\Sigma}\tilde{U}^T \tag{30}$$

For the Laplacian matrix L , we normalize it by:

$$\begin{aligned} L &= D - W \\ &= D^{-\frac{1}{2}}(D-W)D^{-\frac{1}{2}} = I - D^{-\frac{1}{2}} W D^{-\frac{1}{2}} \end{aligned} \tag{31}$$

Where D is the diagonal matrix with $D_{ii} = \sum_{j=1}^{N} W_{ij}$.

Computing the rows sums of $\tilde{W}$:

$$\begin{aligned} \tilde{W} * 1 &= \begin{bmatrix} A & B \\ B^T & B^T A^{-1} B \end{bmatrix} * \begin{bmatrix} 1_l \\ 1_{N-l} \end{bmatrix} \\ &= \begin{bmatrix} A * 1_l + B * 1_{N-l} \\ B^T * 1_l + B^T A^{-1} B * 1_{N-l} \end{bmatrix} \\ &= \begin{bmatrix} a_l + b_l \\ b_{N-l} + B^T A^{-1} b_l \end{bmatrix} \end{aligned} \tag{32}$$

Where a_l, b_l represents the row sums of A and B, b_{N-l} denotes the column sum of B and $\vec{1}$ means a column vector of ones. According to this, we can get:

$$\begin{aligned} D^{-\frac{1}{2}} \tilde{W} D^{-\frac{1}{2}} &= D^{-\frac{1}{2}} \begin{bmatrix} A & B \\ B^T & B^T A^{-1} B \end{bmatrix} D^{-\frac{1}{2}} \\ &= \begin{bmatrix} D_1^{-\frac{1}{2}} A D_1^{-\frac{1}{2}} & D_1^{-\frac{1}{2}} B D_2^{-\frac{1}{2}} \\ D_2^{-\frac{1}{2}} B^T D_1^{-\frac{1}{2}} & D_2^{-\frac{1}{2}} B^T A^{-1} B D_2^{-\frac{1}{2}} \end{bmatrix} \end{aligned} \tag{33}$$

where D_1 is $l * l$ and D_2 is $(N-l) * (N-l)$. For the similarity matrix W, we want to show it can be diagonalized. Supposing that $U = \begin{bmatrix} A & B^T \end{bmatrix}^T A^{-\frac{1}{2}} V \Sigma^{-\frac{1}{2}}$, we can get:

$$\begin{aligned} W &= \begin{bmatrix} A & B^T \end{bmatrix}^T A^{-1} \begin{bmatrix} A & B \end{bmatrix} \\ &= \{\begin{bmatrix} A & B^T \end{bmatrix}^T A^{-\frac{1}{2}} V \Sigma^{-\frac{1}{2}}\} \Sigma \{\Sigma^{-\frac{1}{2}} v^T A^{-\frac{1}{2}} [A \quad B]\} \\ &= U \Sigma U^T \end{aligned} \tag{34}$$

Multiplying from the left by $V\Sigma^{\frac{1}{2}}$ and from the right by $\Sigma^{\frac{1}{2}} V^T$ for the unitary matrix, we can get:

$$\begin{aligned} & V \Sigma V^T \\ &= V \Sigma^{\frac{1}{2}} (U^T U) \Sigma^{\frac{1}{2}} V^T \\ &= V \Sigma^{\frac{1}{2}} (\begin{bmatrix} A \\ B^T \end{bmatrix} A^{-\frac{1}{2}} V \Sigma^{-\frac{1}{2}})^T * \\ & (\begin{bmatrix} A \\ B^T \end{bmatrix} A^{-\frac{1}{2}} V \Sigma^{-\frac{1}{2}}) \Sigma^{\frac{1}{2}} V^T \\ &= V \Sigma^{\frac{1}{2}} (\Sigma^{-\frac{1}{2}} V^T A^{-\frac{1}{2}} [A \quad B]) * \\ & (\begin{bmatrix} A \\ B^T \end{bmatrix} A^{-\frac{1}{2}} V \Sigma^{-\frac{1}{2}}) \Sigma^{\frac{1}{2}} V^T \\ &= A^{-\frac{1}{2}} [A \quad B] \begin{bmatrix} A \\ B^T \end{bmatrix} A^{-\frac{1}{2}} \\ &= A + A^{-\frac{1}{2}} B B^T A^{-\frac{1}{2}} \end{aligned} \tag{35}$$

Because of this property, we reduce the computational complexity and we require only the first k eigenvectors of the Laplacian matrix. Calculating the first k columns of U via

$$U = \begin{bmatrix} \tilde{A} \\ \tilde{B}^T \end{bmatrix} \tilde{A}^{-\frac{1}{2}} (V_N)_{:,1:k} (\Sigma^{-\frac{1}{2}})_{1:k,1;k} \tag{36}$$

Then we normalize U along its rows to get $\tilde{U}$.

$$\tilde{U}_{ij} = \frac{U_{ij}}{\sqrt{\sum_{r=1}^{k} U_{ir}^2}}, i = 1, 2, \ldots, N, j = 1, 2, \ldots, k \tag{37}$$

c. *K-means Step for Normalized Matrix $\tilde{U}$*

Defining $\{u\}_{j=1}^{N}$ is the vectors corresponding to $\tilde{U}'s$ rows. we have extracted cluster numbers k and cluster centers $\{c_1, c_2, c_3, \ldots, c_k\}$ by the method, as introduced in the Section 4. Through the corresponding relationship between normalized matrix $\tilde{U}$ and the clusters centers, we update the cluster centers, when they correspond to same frames. K-means[24] algorithm aims at minimizing the objective function know as squared error function given by:

$$J(\tilde{C}) = \sum_{i=1}^{k} \sum_{\tilde{u}_j \in C_i} ||u_j - \tilde{c}_i||^2 \tag{38}$$

Where $'C_i'$ is the number of data points in i^{th} cluster. According to assign the points to the cluster center whose distance from the cluster center is minimum of all the cluster centers.
Algorithmic steps for k-means: Let $\tilde{U} = \{u_1, u_2, \ldots, u_N\}$ be the elements to be clustered and $\tilde{C} = \{\tilde{c}_1, \tilde{c}_2, \ldots, \tilde{c}_k\}$ be the sets of corresponding centers:

1) Getting k cluster centers $\tilde{C} = \{\tilde{c}_1, \tilde{c}_2, \ldots, \tilde{c}_k\}$;

2) Calculate the distance between each element to be clustered and cluster centers;

3) Assign each element to be clustered to the cluster center whose distance from the cluster center is minimum of all the cluster centers;

4) Recalculate the new cluster center using: $\tilde{c}_i = \frac{1}{C_i} \sum_{j=1}^{C_i} u_i$, Where $'C_i'$ is the number of data points in ith cluster.

5) Recalculate the distance between each element to be clustered and new obtained cluster centers;

6) Stop the iteration until no elements was reassigned, otherwise repeat to step 3).

According to this step and analysis, we can get the N-Frame motion capture data sequences categories information.

D. Behavioral Segmentation Based On Spectral Clustering

We have introduced the method of t-nearest neighbor graphs in Section 5.B. According to this, it can be easily to find that this method is useful to reduce the store of dense similarity matrix. In this way, the dense similarity matrix is represented by the sparse similarity matrix. The algorithm of cluster, which uses the sparse similarity matrix, as follows:

Algorithm 1: clustering uses the sparse similarity matrix.

Input: N-Frame motion capture data $\{f_1, f_2, f_3, \ldots, f_N\}$, every frame is regarded as high-dimensional data point and k: number of desired clusters.

1) Construct similarity matrix $W \in R^{N\times N}$;

2) Modify W to be the sparse matrix S;

3) Compute the Laplcian matrix by $L = I - D^{-\frac{1}{2}} S D^{-\frac{1}{2}}$;

4) Compute the K eigenvectors of L and construct $U \in R^{N\times k}$;

5) Normalize U along its rows to get $\tilde{U}$. $\tilde{U}_{ij} = \frac{U_{ij}}{\sqrt{\sum_{r=1}^{k} U_{ir}^2}}, i = 1, 2, \ldots, N, j = 1, 2, \ldots, k$

6) Use k-means algorithm to cluster N rows of $\tilde{U}$ into k groups, which mean action segments for motion capture data sequences.

In subsection 5.C-b , introducing the Nystrom method, which can be used to cluster for motion capture data sequence.

Algorithm 2: Clustering uses the Nystrom method.

Input: N-Frame motion capture data $\{f_1, f_2, f_3, \ldots, f_N\}$, every frame is regarded as high-dimensional data point; number of samples l, which is determined by 2% the number of motion capture data, k: number of desired clusters and $\{c_1, c_2, c_3, \ldots, c_k\}$: cluster centers, which are determined by the algorithm.

1) Construct $A \in R^{l\times l}$ and $B \in R^{l\times(N-l)}$. The matrix A represent the matrix of similarities between the same points, B be the matrix of affinities between the sample point.

2) Calculate $D = diag(\begin{bmatrix} a_l + b_l \\ b_{(N-l)} + B^T A^{-1} b_l \end{bmatrix})$.

3) Calculate $\tilde{A} = D_{l\times l}^{-\frac{1}{2}} A D_{l\times l}^{-\frac{1}{2}}, \tilde{B} = D_{l\times l}^{-\frac{1}{2}} B D_{(N-l)\times(N-l)}^{-\frac{1}{2}}$.

4) Construct $A + A^{-\frac{1}{2}} B B^T A^{-\frac{1}{2}}$. calculate eigendecomposition $V_N \Sigma_N V_N^T$ for it and ensure the eigenvalues are in decreasing order.

5) Calculate $U = \begin{bmatrix} \tilde{A} \\ \tilde{B}^T \end{bmatrix} \tilde{A}^{-\frac{1}{2}} (V_N)_{:,1:k} (\Sigma^{-\frac{1}{2}})_{1:k,1;k}$ as the first k eigenvector of the Laplacian matrix.

6) Normalize U along its rows to get $\tilde{U}$.$\tilde{U}_{ij} = \frac{U_{ij}}{\sqrt{\sum_{r=1}^{k} U_{ir}^2}}, i = 1, 2, \ldots, N, j = 1, 2, \ldots, k$

7) Use k-means to cluster N rows of $\tilde{U}$ into k classes.

Two methods of these can realize clustering motion capture data sequence into motion fragments.

VI. REFINE BEHAVIORIAL SEGMENTATION RESULTS

According to temporal reverting, we can get $S = \{s_1, s_2, s_3, \ldots, s_{k'}\}$ for N-Frame motion capture data sequence, where s_i represents the subsequence after clustering and temporal reverting. Setting the single representation $G = \{g_{c_i}\}_{c\in\{c_1,c_2,\ldots,c_k\},\mathrm{i}=1,2,\ldots,N}$. g_{c_i}=1 if f_i belongs to class c, otherwise g_{c_i}=0.

According to extensive experiments, we can find that clustering and temporal reverting for motion capture data can lead some error. We define energy function to reduce noise and realize behavior segmentation. Because of the temporal property, behavior segmentation points locate in adjacent motion subsequences. For each motion subsequence, we only calculate it for belonging to the former one or the last one. The energy function is defined by:

$$F(S, G) = \sum_{i,j\in 1,2,\ldots,k'} \sum_{c_i=1}^{k} g_{c_i} dist(s_i, s_j) \tag{39}$$

Where s_i represents motion subsequences for motion capture data after clustering and temporal reverting. Pavel Senin [25] provided the good resolution for calculating the distance for temporal sequences, which is named dynamic time warp (DTW).

$$\begin{aligned} DTW(s_i, s_j) &= dist(s_i, s_j) \\ &= \arg\min\{\sum_{f_{i'}\in s_i, f_{j'}\in s_j} ||f_{i'} - f_{j'}||\} \end{aligned} \tag{40}$$

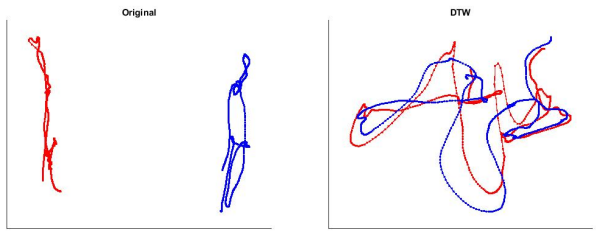
Dynamic programming is applied to deal with behavior segmentation. The formula is defined as follows:

$$J(S^{'}) = \min\{J(S^{'} - 1) + min\{F(S^{'} - 1, G)\}\} \tag{41}$$

In this way, we can realize behavior segmentation. After clustering, long action subsequences can be considered as independent behavior and we use the energy function to deal with error subsequences. Where s_i represents action subsequences for motion capture data after clustering and temporal reverting. The Figure 3 shows an example of two motion capture data sequences aligned by DTW.

VII. EXPERIMENTS

In order to prove the feasibility and effectiveness of this method, we use the Carnegie Mellon University motion capture database [CMU][26], whose data were captured with a Vicon optical motion capture system of 12 MX-40 cameras at 120HZ. Extensive experiments are conducted on the dataset of multi-behavior



(a) (b)

Figure 3: The result of two motion capture data sequences aligned by DTW, which is demonstrated in 2-dimensional space. (a) represents original sequences and (b) shows aligned result.

motion capture data from CMU database. each motion capture data sequence is the combination of roughly several natural actions (walking, running, punching, jumping and so on.).

The greatest advantage of our method for behavior segmentation is automatic and without human intervention. For arbitrary motion capture data, which contains several behaviors, we can realize behavior segmentation. Figure 4 shows the segmentation results obtained through PPCA, GMM, our methods (t-nearest and Nystrom method) and manual segmentation. In every chart, x-axis represents the motion frames and y-axis defines methods abbreviation. Each chart contains five small bars. From the top downwards, each bar represents the results of behavior segmentation by PPCA, GMM, t-nearest neighbors to cluster, Nystrom method applied to cluster and manual segmentation. The first four methods use black vertical line to represent the results of segmentation.

For the manual segmentation, it use black strip to represent segmentation results. Since the continuity of behaviors and the restriction of human recognition, the human segmentation labels are not single frame and they are a short sequence of frames. Manual segmentation results are regarded as ground truth for segmentation of motion capture data. Due to the continuity of motion, we think that segmentation results located in the vicinity of ground truth are segmentation successful.

According to experiments, we can find that our method not only can segment motion capture data, but also can mark the same behavior segments. We use same color to mark the same behavior segments. The experiments results are close to ground truth. That as long as the segmentation results falling in this sequence are the right points of segmentation. We compare these behavior segmentation algorithms in the standard precision/recall framework.

$$\begin{cases} precision = \frac{\#reportedcorcuts}{\#reportedcuts} \times 100\% \\ recall = \frac{\#reportedcorcuts}{\#corcuts} \times 100\% \end{cases} \tag{42}$$

where #reportedcorcuts indicates the reported correct cuts, #reportedcuts is the total number of reported cuts and #corcuts represents the total correct cuts. The closer precision and recall are to one, the more effective the algorithm is. Table 2 gives scores of precision and recall for PPCA, GMM, t-nearest and Nystrom method.

Table 2: Precision and recall scores for the PPCA,GMM,t-nearest and Nystrom method

Behavior Methods	Precision	Recall
PPCA	80.83%	87.38%
GMM	54.00%	72.97%
t-nearest	90.09%	90.91%
Nystrom	92.79%	93.63%

VIII. DISCUSSION AND CONCLUSION

In this paper, we introduce a novel automatically method to segment motion capture data. Firstly, we use a novel method to extract cluster numbers k, which can be regarded as the number of behavior contained, and cluster centers. This cluster algorithm is based on local density for input data. Then, we construct undirected weighted graph, which uses motion capture data. In order to deal with this NP hard problem, we analyze relationship between graph cut and spectral clustering. According to this, the problem of behavior segmentation for arbitrary motion capture data sequence is converted to spectral clustering problem. According to the priori knowledge of behavior numbers and cluster centers, the application of t-nearest neighbors and Nystrom method used respectively to cluster for motion capture data sequence. Defining energy function to refine segmentation for motion capture data. These method can reduce computational complexity, save restore space and improve calculation speed. Energy function, which is defined by us, can reduce the error of segmentation results.

Although we have reduce the restore space for this method, it takes up remarkable restore space. In future work, we should further improve the calculate speed and reduce restore space. Meanwhile, to further study the effect of single behavior fragments on the motion analysis and motion synthesis.

ACKNOWLEDGMENTS

This research is partially supported by National Natural Science Foundation of China (No. 61370127, No.61473031, No.61472030), Program for New Century Excellent Talents in University (NCET-13-0659), Fundamental Research Funds for the Central Universities (2014JBZ004), Beijing Higher Education Young Elite Teacher Project (YETP0583). The opinions expressed are solely those of the authors and not the sponsors.

References

[1] Paulo Sousa, João L Oliveira, Luis Paulo Reis, and Fabien Gouyon. Humanized robot dancing: humanoid motion retargeting based in a metrical representation of human dance styles. In *Progress in Artificial Intelligence*, pages 392–406. Springer, 2011.

[2] Jianyuan Min and Jinxiang Chai. Motion graphs++: a compact generative model for semantic motion analysis and synthesis. *ACM Transactions on Graphics (TOG)*, 31(6):153, 2012.

[3] Adso Fern'ndez-Baena, Antonio Susin, and Xavier Lligadas. Biomechanical validation of upper-body and lower-body

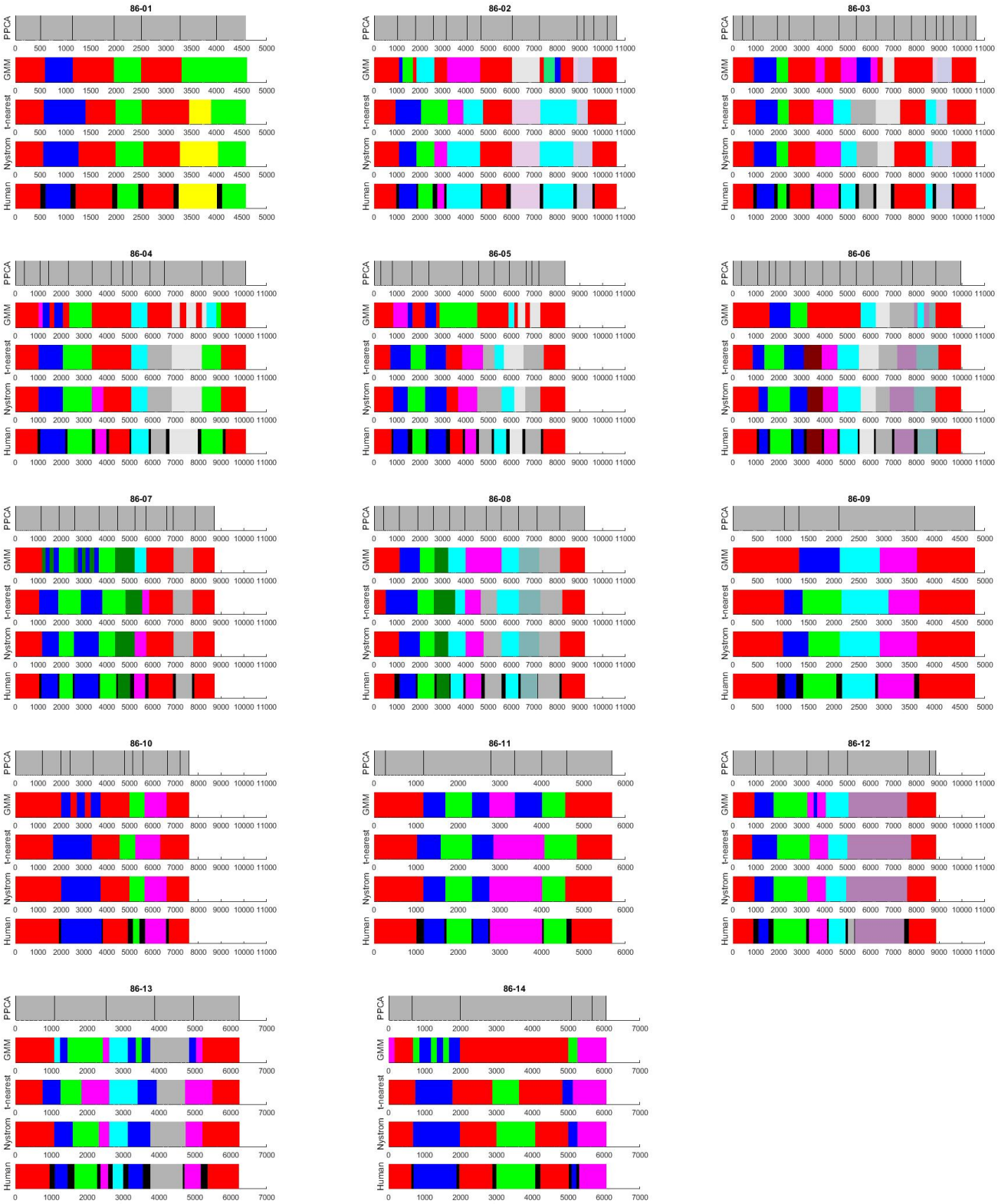


Figure 4: The results of segmentation. It shows the segmentation results obtained through PPCA, GMM, t-nearest neighbors to cluster, Nystrom method applied to cluster and manual segmentation. Different behaviors are indicated by different colors. The five colors stripes in the same chart represent the same motion capture data sequence.

joint movements of kinect motion capture data for rehabilitation treatments. In *Intelligent Networking and Collaborative Systems (INCoS), 2012 4th International Conference on*, pages 656–661. IEEE, 2012.

[4] Yao-Jen Chang, Shu-Fang Chen, and Jun-Da Huang. A kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities. *Research in developmental disabilities*, 32(6):2566–2570, 2011.

[5] Ilaria Pasciuto, Sergio Ausejo, Juan Tomás Celigüeta, Ángel Suescun, and Aitor Cazón. A hybrid dynamic motion prediction method for multibody digital human models based on a motion database and motion knowledge. *Multibody System Dynamics*, 32(1):27–53, 2014.

[6] Paul Fearnhead. Exact and efficient bayesian inference for multiple changepoint problems. *Statistics and computing*, 16(2):203–213, 2006.

[7] Eamonn Keogh, Selina Chu, David Hart, and Michael Pazzani. An online algorithm for segmenting time series. In *Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on*, pages 289–296. IEEE, 2001.

[8] Xiang Xuan and Kevin Murphy. Modeling changing dependency structure in multivariate time series. In *Proceedings of the 24th international conference on Machine learning*, pages 1055–1062. ACM, 2007.

[9] Meinard Müller, Andreas Baak, and Hans-Peter Seidel. Efficient and robust annotation of motion capture data. In *Proceedings of the 2009 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pages 17–26. ACM, 2009.

[10] Fengjun Lv and Ramakant Nevatia. Recognition and segmentation of 3-d human action using hmm and multi-class adaboost. In *Computer Vision–ECCV 2006*, pages 359–372. Springer, 2006.

[11] Okan Arikan, David A Forsyth, and James F O'Brien. Motion synthesis from annotations. *ACM Transactions on Graphics (TOG)*, 22(3):402–408, 2003.

[12] Feng Zhou, F Torre, and Jessica K Hodgins. Aligned cluster analysis for temporal segmentation of human motion. In *Automatic Face & Gesture Recognition, 2008. FG'08. 8th IEEE International Conference on*, pages 1–7. IEEE, 2008.

[13] Feng Zhou, F Torre, and Jessica K Hodgins. Hierarchical aligned cluster analysis for temporal clustering of human motion. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(3):582–596, 2013.

[14] Michal Balazia, Jan Sedmidubsky, and Pavel Zezula. Semantically consistent human motion segmentation. In *Database and Expert Systems Applications*, pages 423–437. Springer, 2014.

[15] J. Barbič, A. Safonova, JY. Pan, C. Faloutsos, JK. Hodgins, and NS. Pollard. Segmenting motion capture data into distinct behaviors. In *Proceedings of Graphics Interface 2004*, pages 185–194. Canadian Human-Computer Communications Society, 2004.

[16] Jehee. Lee, Jinxiang. Chai, Paul S.A. Reitsma, Jessica K. Hodgins, and Nancy S. Pollard. Interactive control of avatars animated with human motion data. In *ACM Transactions on Graphics (TOG)*, volume 21, pages 491–500. ACM, 2002.

[17] Jing Wang and Bobby Bodenheimer. An evaluation of a cost metric for selecting transitions between motion segments. In *Proceedings of the 2003 ACM SIGGRAPH/Eurographics symposium on Computer animation*, pages 232–238. Eurographics Association, 2003.

[18] J Macqueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, 1967*.

[19] Alex Rodriguez and Alessandro Laio. Clustering by fast search and find of density peaks. *Science*, 344(6191):1492–1496, 2014.

[20] Jinhui Yuan, Huiyi Wang, Lan Xiao, Wujie Zheng, Jianmin Li, Fuzong Lin, and Bo Zhang. A formal study of shot boundary detection. *Circuits and Systems for Video Technology, IEEE Transactions on*, 17(2):168–186, 2007.

[21] Lars Hagen and Andrew B Kahng. New spectral methods for ratio cut partitioning and clustering. *Computer-aided design of integrated circuits and systems, ieee transactions on*, 11(9):1074–1085, 1992.

[22] Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(8):888–905, 2000.

[23] Wen-Yen Chen, Yangqiu Song, Hongjie Bai, Chih-Jen Lin, and Edward Y Chang. Parallel spectral clustering in distributed systems. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(3):568–586, 2011.

[24] Tapas Kanungo, David M Mount, Nathan S Netanyahu, and Angela Y Piatko. An efficient k-means clustering algorithm: Analysis and implementation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7):881–892, 2002.

[25] Pavel Senin. Dynamic time warping algorithm review. *Information and Computer Science Department University of Hawaii at Manoa Honolulu, USA*, pages 1–23, 2008.

[26] CMU. Carnegie mellon university graphics lab: Motion capture database. `http://motioncapture.cs.cmu.edu`.

Richard: Towards a Dialogue System Supporting Automatic Event Identification

Yibin Jiang
RWTH Aachen
Templergraben 55
52062 Aachen, Germany
yibin.jiang@rwth-aachen.de

Tiansi Dong, Armin B. Cremers
B-IT University of Bonn
Dahlmannstraße 2
53113 Bonn, Germany

Joachim Köhler
Fraunhofer IAIS
Schloss Birlinghoven
53754 St. Augustin, Germany

Abstract

A dialogue system, Richard, for free communication on daily issues is under construction. Daily issues are classified into tree-structured event knowledge-base. For each event topic, we retrieve dialogue examples from the Internet. Basic-level Upper Ontologies (BLUO) are proposed to categorize words of these examples. Event identification is realized by comparing the similarity between BLUOs of an input sentence and BLUOs of example sentences in the event knowledge-base. Cosine similarity algorithm is adopted to select the most possible event domain. BLUOs of an input sentence are combined to the sentence ID. A difference distance acts as a metric to determine the most similar ID in the script with input sentence ID. An example of free dialogue within Greeting-Shopping-Restaurant domains is described. On-going work on meaning-based dialogue system and dialogue-based learning is outlined.

1. Introduction

"A dialogue system is a computer program that communicates with a human user in a natural way" Arora et al. (2013). Dialogue systems have many application areas, e.g. call-center service, shopping assistance, traveling service. Research on human-computer dialogue system can be dated back to Eliza by Weizenbaum (1966). The first dialogue system Eliza was based-on word-level pattern matching method, e.g, Weizenbaum (1966), Norvig (1992). Eliza can act as a psychotherapist to soothe a human patient. Although Eliza system also relies on the context, its strategy is not sufficient for carrying out daily dialogues since people often talk from one context to other context.

A recent milestone would be IBM's Watson, High (2012), which is based-on statistical reasoning supported by a large knowledge-base. Watson is able to answer quiz questions in English. LogAnswer is a German question-answer system based on formal logic reasoning supported by the German Wikipedia, Furbach et al. (2010), Dong et al. (2011). For human-computer dialogue in everyday life, the dialogue system might not need a very large knowledge base, however, it shall be aware of the dynamic chances of topics, and adjust its language analysis strategies. Another difference from pure question-answer systems is that a dialogue system sometimes should also raise questions, make suggestions. This is, question answering systems are passive, while dialogue systems can be active. We set out two criteria for dialogue systems for the daily-use as follow: (1) they should be able to automatically identify communication domains; (2) they should be able to accept ungrammatical utterances.

We report an approach to identifying dialogue topics and switching among dialogue domains. Section 2 introduces the basic idea for domain identification; Section 3 describes the tree-structured domain knowledge system; Section 4 presents the cosine similarity algorithm for automatic domain identification; Section 5 elaborates the usage of BLUOs for sentence ID identification in the script; Section 6 describes the Richard dialogue system; Section 7 lists two pieces of on-going work: meaning-based domain identification, and dialogue-based learning of domain knowledge.

2. The research on "Basic Level Upper Ontologies" in daily dialogues

Rosch et al. (1976) identified "basic level category" as a certain level of taxonomies of objects, which *carry the most information and the highest category cue validity*. Basic level categories are *the first categorizations which people make during perception of the environment*, and also are *the earliest categories named by children*. What we inspired is that when people change environments, e.g. going out of home, walking in subways, looking around in shopping malls, basic-level words used in their communications would be a valid cue to identify the changes of environments. We carried out a computational research to identify basic-level words and their upper layer ontologies in daily

DOI reference number: 10.18293/DMS2016-020

dialogues, which we name them "Basic Level Upper Ontologies".

We collect dialogues from Internet, which have been classified under functional domains, such as "daily greeting", "shopping". For each domain, we manually go through these dialogues and identify basic-level categories, and construct a word-category dictionary. For the basic-level categories, we identify their upper ontologies, which serve as certain mid-level taxonomies between basic-level categories and the domain. When a new utterance is received, the system will first map the words in the utterance into basic level categories, and reasoning on the possible domain.

So, Basic Level Upper Ontologies are used for (1) normalization of input utterances, (2) identification of dialogue domain, e.g. topics, environment, based on the knowledge-base, and (3) identification of sentence ID in the script. For the utterance "I go to Italian restaurant", we would map "I" into the category EgoPerson, "Italian" into Nationality, and "restaurant" into FoodPlacePublic. The category name is constructed as follows: an upper category of the word and several characteristic features are combined and ordered alphabetically: "I" is a Person with the feature Ego, "restaurant" is a Place with feature of Food and Public.

3. Knowledge of Events and Domains

In the Richard system, domains are understood as families of related event structures. Dialogues start with greeting and end with separating. We introduce the greeting domain as the first domain, and the good-bye domain as the last domain. Domains can be overlapped. For example, shopping domains and restaurant domains may share payment domain in common. Events are spatial-temporal entities, e.g., Zacks and Tversky (2001). Two events can follow 13 qualitative temporal relations, as described in Allen (1983): for example, a greeting event can be directly followed by a going-to-cafeteria event.

Domains are structured in a lattice. Every domain in the lattice is a node. We use a child of a node to represent the sub-domain of a domain. All nodes or leaves with the same parent have the parallel relation. If two domains have the overlapping part, we extract this part from them and put it as a new domain in the subtrees of both domains.

For Richard system, we introduced restaurant domain and shopping domain. As to restaurant domain, it contains before-restaurant, in-restaurant and after-restaurant domains. Under the in-restaurant domain, there are two parallel domains, order domain and payment domain. After-restaurant domain has positive domain which means positive comments on the restaurant and negative domain which means negative comments. As for the shopping domain, it contains before-shopping, in-shopping and after-shopping domains. There are look domain and payment domain under the in-shopping domain. The structure of domains in Richard system is shown in Figure 1.

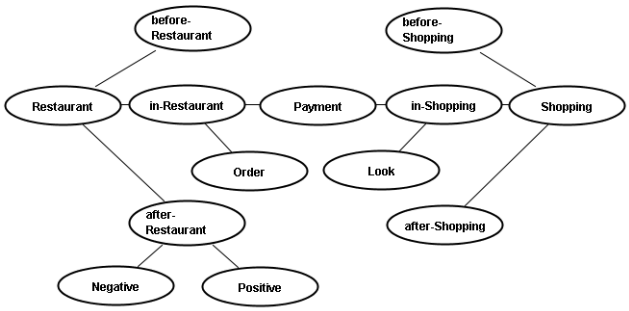


Figure 1. The Sample of Domain Structure

4. Automatic Event Identification

The aim of the research is to automatically identify event domain. This is carried out by a deviated cosine similarity function, followed by a decision making process as follows: for each domain, we define a domain ID vector, whose size is the number of all BLUOs within this domain appeared in dialogue corpus, and each vector element is corresponding to a BLUO and set to frequency number of this BLUO in the corpus. In the current dialogue system, the restaurant domain ID vector has the rank of 22, the shopping domain ID vector has the rank of 19. Given an input sentence, we define the input vector as follows: it has the same rank of the domain ID vector; if there is a word in the input sentence belongs to a BLUO in the domain, the corresponding element of the input vector is set to 1, otherwise 0. The input sentence is classified into the domain with highest cosine similarity value. If there are more than two maximum cosine values, among which there is the domain to which the last input sentence belong, we choose the last domain, otherwise, we randomly choose a dialogue domain.

Just an example, suppose that the restaurant domain ID vector has the rank of 3, which corresponds to three BLUOs: Restaurant, Eat, and Hungry, each appears 3, 4, 2 times in the dialogue corpus. The restaurant domain ID vector turns out to be $[3, 4, 2]$; suppose that the input sentence is "I want to eat", among those words, "I" belongs to the BLUO of EgoPerson, "want" is neglected, "eat" belongs to Eat. So, the input vector is $[0, 1, 0]$. The cosine similarity is $\frac{3*0+4*1+2*0}{\sqrt{3^2+4^2+2^2}\sqrt{0^2+1^2+0^2}} = 0.7428$. As to the shopping domain, if its ID vector has the rank of 3 as well. The corresponding BLUOs are Shop, Buy, and trolley which appear 3, 3, 1 times respectively in the dialogue

corpus. The shopping domain ID vector is $[3, 3, 1]$. However, in the above sample sentence, "I want to eat", none of the shopping domain BLUOs appear so the input vector is $[0, 0, 0]$. Since in the cosine similarity, the denominator is $\sqrt{3^2 + 3^2 + 1^2}\sqrt{0^2 + 0^2 + 0^2} = 0$, the cosine similarity is assigned to 0. As 0.7428 is greater than 0, the sample sentence, "I want to eat", is regarded as belonging to the restaurant domain. The algorithm for event domain identification is outlined as follows.

```
d = current domain;
while input not empty do
    get BLUOs and their frequencies;
    for number of domains do
        construct input vector based on domain
        BLUOs;
    end
    calculate correlations by similarity;
    list ← get biggest correlations;
    if d not in list then
        d = random(list);
    else
    end
end
```

Algorithm 1: Event Domain Identification Algorithm

5. Sentence ID Identification

All BLUOs contained in an input sentence can be combined together as a sentence ID used to choose the response. BLUOs in the sentence ID is reordered alphabetically. In the script, a response is categorized according to an ID of a sentence which that response responds to. We assign a difference distance to different sentence IDs. The difference distance between two IDs are the sum of distinct BLUOs in both IDs. Assume $id1$ and $id2$ are two sentence IDs and BLUO() is a BLUO extraction operator. The formula of the difference distance is below.

$$D(id1, id2) = BLUO(id1) \vee BLUO(id2) - BLUO(id1) \wedge BLUO(id2)$$

For example, the difference distance between FoodOrder and FoodQuantity is 2. If a sentence ID $id1$ has all BLUOs contained in another sentence ID $id2$, we call $id2$ is the super-set of $id1$ and $id1$ is the subset of $id2$. Food is a subset of FoodQuantity.

In Richard system, we search in a bidirectional way to find one smallest super-set called down-to-up search and one biggest subset called up-to-down search of the input sentence ID in the script. Then, we calculate the difference distance between the input sentence ID and its super-set and between the input sentence ID and its subset. We choose the super-set or subset based on the smaller difference distance. For example, in the restaurant domain, we have two IDs FoodDirectionLocationState and Location. The id from the input sentence is DirectionLocation. The difference distance between DirectionLocation and FoodDirectionLocationState is

$$BLUO(DirectionLocation) \vee BLUO(FoodDirectionLocationState) - BLUO(DirectionLocation) \wedge BLUO(FoodDirectionLocationState) = 2$$

The difference distance between DirectionLocation and Location is

$$BLUO(DirectionLocation) \vee BLUO(Location) - BLUO(DirectionLocation) \wedge BLUO(Location) = 1$$

Since the difference distance between DirectionLocation and Location is smaller than that between DirectionLocation and FoodDirectionLocationState, we choose the rules below the ID Location.

```
cd = chosen domain;
tag = BLUOsTag(input);
tag1 = up2downSearch(tag, cd);
tag2 = down2upSearch(tag, cd);
if tag1.length < tag2.length then
    return tag1;
else
    return tag2;
end
```

Algorithm 2: Sentence ID Identification Algorithm

6. System Implementation

Sample dialogue sentences are obtained from the Internet. For the presented work, we used sample dialogues at `http://www.englishspeak.com/de/english-lessons.cfm` and `http://www.eslfast.com/robot/`. The model of Richard contains five main components as in Figure 2. Inside Richard, there are input processor, domain identifier,

sentence ID identifier and rule selector. Outside the system, Richard retrieves the rules from a script. The circle handle is a provided interface and the semicircle handle is a required interface.

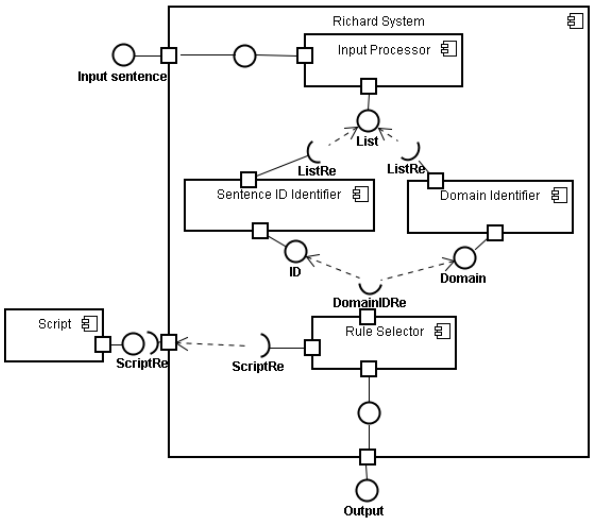


Figure 2. The Model of Richard

We divide the construction of Richard system into two major parts. They are the knowledge-based part and the dialogue part. In the knowledge-based part, we retrieve sample dialogue sentences from the Internet. Based on domains we choose, these sentences are categorized to each of them. Then, we extract the BLUOs from each sentence and construct the domain ID vector for each domain based on the BLUOs of this domain. Finally, we assign each sentence under an ID of a sentence which the sentence assigned by us responds to.

In the dialogue part, when a sentence comes into the input processor, it will be divided into a word list and each word is replaced with its BLUO. This forms a BLUO list sent to both domain and sentence ID identifiers. In the domain identifier, for each domain ID vector, the BLUO list is transformed to corresponding input vector and the cosine similarity between these two vectors are calculated. The domain with the biggest similarity is chosen. In the sentence ID identifier, the BLUO list is formed to a sentence ID. Based on this ID, an up-to-down search finds the biggest subset and a down-to-up search finds the smallest superset. The difference distance calculation is conducted for the most similar ID in the script. At last, a rule selector selects a response under this ID in the chosen domain randomly.

7. On-going Work

Based on Richard system, we have two more on-going work. Firstly, our dialogue system is based on the Basic Level Upper Ontologies defined manually. Human use language to communicate feelings and ideas. A model of ideas can be formed based on the semantic approach Simmons (1973). Therefore, we want to switch the domain in terms of meaning called meaning-based domain identification. The second task is to make the system learn the domain knowledge, which is dialogue-based learning of domain knowledge.

Acknowledgment

The work reported here was carried out during the “IPEC Winter School 2015 on Speech Technology and Python” from Feb. 2, 2015 to Mar. 31, 2015 at B-IT.

Sample dialogues with Richard

```
HUMAN>Hi, Richard.
[Domain:GreetingDomain]
RICHARD>Hey
HUMAN>How are you?
[Domain:GreetingDomain]
RICHARD>Fine, how are you doing?
HUMAN>Fine, too.
[Domain:GreetingDomain]
RICHARD>What do you want to do?
HUMAN>I am hungry. I want to eat.
[Domain:BeforeRestaurantDomain]
RICHARD>Let’s go to an Italian
    restaurant now.
HUMAN>How about this restaurant?
[Domain:OrderDomain]
RICHARD>Yeah, it is. I come here
    all the time.
HUMAN> What do you want to order?
ANALYSIS> [Domain:OrderDomain]
RICHARD> I’m not sure. I haven’t
    decided yet. Can you recommend
    something?
HUMAN> Fish.
ANALYSIS> [Domain:OrderDomain]
RICHARD> Good choice.
HUMAN> Let’s pay now.
ANALYSIS> [Domain:PaymentDomain]
RICHARD> Yes, please.
HUMAN>I want to go shopping.
[Domain:BeforeShoppingDomain]
RICHARD>What time are you leaving?
HUMAN>I will leave now.
[Domain:BeforeShoppingDomain]
RICHARD>Will you buy a ham
    sandwich for me at the store?
HUMAN>Ok.
```

```
[Domain:BeforeShoppingDomain]
RICHARD>But I have no money now.
HUMAN>You can give me later.
[Domain:BeforeShoppingDomain]
RICHARD>Thanks.
HUMAN> I bought a lot of food.
ANALYSIS> [Domain:AfterShoppingDomain]
RICHARD> Where did you buy it?
HUMAN> In the supermarket.
ANALYSIS> [Domain:AfterShoppingDomain]
RICHARD> Was it expensive?
HUMAN> They are cheap
ANALYSIS> [Domain:AfterShoppingDomain]
RICHARD> Really?
HUMAN>[Abort]
Bye! See you next time!
```

References

Allen, J. F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843.

Arora, S., Batra, K., and Singh, S. (2013). Dialogue System: A Brief Review. pages 2–5.

Dong, T., Furbach, U., Glöckner, I., and Pelzer, B. (2011). A Natural Language Question Answering System as a Participant in Human Q&A Portals. In *IJCAI*, pages 2430–2435, Barcelona, Spain.

Furbach, U., Glöckner, I., and Pelzer, B. (2010). An application of automated reasoning in natural language question answering. *AI Communication*, 23(2-3):241–265.

High, R. (2012). The Era of Cognitive Systems: An Inside Look at IBM Watson and How it Works. page 14.

Norvig, P. (1992). Paradigms of Artificial Intelligence Programming: Case Studies in Common Lisp. pages 151–174.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., and Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8:382–439.

Simmons, R. F. (1973). Semantic networks : Their computation and use for understanding english sentences. *Computational Models of Thought and Language*, pages 63–113.

Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9:36–45.

Zacks, J. M. and Tversky, B. (2001). Event structure in perception and conception. *Psychological bulletin*, 127(1):3–21.

A Tool for the Semantic Analysis and Recommendation of Videos in e-Learning

Mauro Coccoli and Gianni Vercelli
DIBRIS
University of Genoa
Genoa, Italy
{mauro.coccoli, gianni.vercelli}@unige.it

***Abstract*—Video lessons are increasingly adopted in education, especially in universities and lifelong learning projects. Their popularity is due to the people's familiarity with video and to other intrinsic characteristics of this medium, such as the message rapidity and its reproducibility. Accordingly, Massive Open Online Courses are gaining a prominent role in both formal and informal education and many universities provide video courses for their students through suited platforms or even freely accessible to everyone. To improve the effectiveness of video lessons and to make them part of a wider learning environment, we decided to investigate the possibility of making a system that, starting from a video, can suggest further "readings". Such a system is thought for independent lifelong learners, for regular students, for teachers and instructional designers as well.**

Keywords-Cognitive computing; video lessons; MOOCs; technology enhanced learning; lifelong learning.

I. INTRODUCTION

The use of video lessons is becoming more and more popular among students for a variety of reasons. Since the raise of specialized portals such as, e.g., Vimeo and YouTube, a huge number of tutorials and *howtos* were made available spontaneously by the users in the form of video, which cover almost any topic, ranging from, e.g., makeup to housekeeping and cars or bikes maintenance. This new wave of User Generated Contents (UGC) dramatically changed the use of the Web, the Internet traffic profiles, and the approach of people towards learning. In fact, people started looking for videos instead of text to clarify concepts, see examples, find additional information on the topic they are interested in or are studying. Yet, there is a lack of organization in these gigantic, general purpose video repositories. Especially, it is difficult to obtain information on the selected resources and get details on their content, hence accessing relevant video fragments is a hard task. Considering the fact that YouTube is worldwide the second largest source of information on the Internet after Google [1], the lack of a correct cognitive alignment between the available videos and the expected learning outcomes with respect to the given audience is still an open problem. To cope with this problem, a semantic support would be required, to enable the effective use of metadata and to improve research capabilities. Unfortunately, in most cases, videos that can be found on the Internet are user-generated and self-produced and we cannot expect a consistent behavior from such a very large, heterogeneous group of persons. Then, alternative solutions must be found so that the needed information can be extracted from videos through automated tools and sophisticated mining techniques.

It is worthwhile noticing that teachers are directly involved in this revolution both as actors and directors. Many of them are producing video courses for enhancing their traditional classes and/or for creating new educational projects in MOOCs (Massive Open Online Courses) platforms, whose number is fast-growing. In fact, following the wave of the OpenCourseWare initiative launched by the Massachusetts Institute of Technology (MIT) and the success of services such as, e.g., Coursera, many universities decided to counteract this phenomenon and provide own video courses through internal and national portals. It is the case of national French platform FUN (*France Université Numérique*), based on the edX platform, and of the recently released Eduopen initiative in Italy, which sees the authors of this paper directly involved in the production of video lessons for the "Computer Science" modules.

Beside their popularity, video lessons are to be considered an effective also from the learning point of view, for they offer the following advantages:

(i) enhance engagement,

(ii) improve communication,

(iii) improve clarity,

(iv) show examples,

(v) offer the possibility of using captions and exploiting videos recorded in different languages,

(vi) introduce interactivity elements,

(vii) make educational resources freely accessible with reduced costs for development and delivery [2].

Given these premises, we forecast a near future in which it will be easier and easier to find videos on the Web covering a wide variety of topics, also in academic fields, that can be considered reliable due the prestige of the authors and to the presence of metadata qualifying their content and authors as well. Hence, in this paper, we showcase a demonstrative prototype recently developed, which was designed to perform the following activities:

DOI reference number: 10.18293/DMS2016-018

(i) find interesting videos over the Web on a given matter,

(ii) play videos,

(iii) make a real-time transcript of the video, which can be used for other activities such as collecting text for annotation and semi-automatic translation from foreign language to individuals' mother language,

(iv) process the resulting text with advanced semantic algorithms to extract concepts and relevant information and to find related material in the form of more videos, additional readings, link to other learning assets.

We highlight that the presented prototype is based on the use of cognitive computing solutions and a cloud infrastructure, whose technical details are depicted in the following.

The remainder of the paper is organized as follows. Firstly, in Section 2, an overview of the reasons for using videos in education is given, also covering issues related to the adoption of MOOCs. Section 3 describes new opportunities offered by cognitive computing and how they can be profitably exploited in education. Accordingly, Section 4 provides a detailed description of the proposed system that aims to be a valuable support to both teachers and learners. Finally, Section 5 summarizes findings and gives a glance on future works.

II. Using Videos in e-Learning and MOOCs

In the last five years, MOOCs have emerged as a worldwide adopted, low-cost e-learning technology to deliver video-based educational contents on a wide variety of topics. Given the lack of entry barriers (prerequisites), a lot of universities offers platforms and facilities for their students to enroll in MOOCs. However, only a small fraction of enrolled students actively participates in the various activities, such as viewing video lectures, completing quizzes and homework, and so on. In this respect, the dropout phenomenon has been studied [3], [4]. Several researchers have analyzed usage logs to determine the factors associated with MOOCs' high attrition rate and to develop methods that predict how participants will perform and whether they are likely to drop out [5]. This study shows the results of a research focused on the perception of students about learning in a cooperative MOOC course. It is a university MOOC course, since it is inserted into the formal teaching of Master Degree in Primary Education, although open to participation by all interested persons. The perception of students is evaluated through the questionnaire Technology Acceptance Model (TAM) adapted to the context of learning in a massive open online course. Three variables associated with the perception of students on learning in the course are discussed:

(i) the academic level of participants (undergraduate *vs.* non-degree students),

(ii) the type of participation in the course (hidden, moderately hidden, active, individualistic and collaborator),

(iii) the gender variable.

Although the perception of the students in income, motivation and ease of use of a MOOC course is high, there are some differences between the variables analyzed. It is noted that no grade students perceive more positively the methodology with which one has to work in a MOOC. However, the type of participation does not significantly influence the perception of students about learning in a MOOC course. In addition, males scored significantly higher on both the motivation to learn through a MOOC, and the perception of its usefulness for learning. The results show good acceptance by students and the use of massive open online courses in regulated university contexts, and suggest a greater acceptance in learning environments linked to lifelong learning and professional development.

III. Semantics and Cognitive Computing

The above-cited tool, whose details will be duly discussed in Section 4, relies on sophisticated cognitive computing techniques made available through the rich set of services and APIs accessible within the IBM Bluemix platform [6]. Owing to the IBM Academic Initiative, universities can freely experiment Watson and other functionalities. In particular, the platform provides programmers with a plenty of services based on the Watson artificial intelligence [7] and the DeepQ&A algorithm, enabling the creation of complex applications and, thus, giving a significant contribution to the increasing success of cognitive computing applications in many fields of applications, including university classes [8]. In more detail, Watson is a full-featured cognitive computing system for the research and development of cognitive systems and services [9], [10], which can easily interoperate with other applications and legacy systems as well. It is a valuable commercial tool, exposing a varied catalog of building blocks, which allow rapid prototyping of applications with advanced features. For example, among the others, we cite the following boilerplates[1] that can help enhance, scale, and accelerate human expertise in the direction of making software applications smarter and smarter:

(i) Concept Expansion - performs text analysis and can learn similar terms as well as words or phrases, based on context. Such a tool enables users to rapidly create a lexicon and a set of related terms from data sets of text fragments or collections of documents. Then, the output can be used to provide further understanding of data and improve text analytics pipelines;

(ii) Concept Insights - looks for associations of concepts inside sets of documents provided by users with a pre-existing graph of concepts based on the renowned free encyclopedia Wikipedia. Accordingly, the service identifies links of two types: explicit links in the case a document directly mentions a concept, and implicit links, which connect the input documents to relevant concepts that are not directly mentioned. This service can also search for documents that are relevant to a concept or collection of concepts by exploring both the explicit and the implicit links;

(iii) Dialog - allows developers designing the interaction mechanisms of an application with an end-user, based on a

[1] Borrowed from https://console.ng.bluemix.net/catalog - Can be subject to modifications.

natural conversational interface. In practice, this service enables computer applications to use natural language and this capability can be profitably exploited in a variety of situations such as, e.g., automatically respond to user questions, walk users through processes or applications, or even hand-hold users through difficult tasks;

(iv) Natural Language Classifier - applies cognitive computing techniques to analyze sentences or phrases in a given corpus and return the best matching classes;

(v) Relationship Extraction - parses sentences into their various components, looking for relationships between the components;

(vi) Speech-to-text and Text-to-speech - the former converts the human voice into the written word and the latter processes text and natural language to generate synthesized audio output complete with appropriate cadence and intonation even if on a limited set of languages;

(vii) Tone Analyzer - leverages cognitive linguistic analysis to identify tones that people show in their languages;

(viii) Visual Recognition - allows analyzing the visual appearance of images or video frames to understand what is happening in a scene. Based on machine learning technology, the semantic classifiers recognize many visual entities, such as, e.g., settings, objects, and events;

(ix) Tradeoff Analytics - helps people make better choices while taking into account multiple, often conflicting, goals that matter when making that choice.

Owing to such an easy-to-use arsenal of services with stunning functionalities, new opportunities arise for both designers and programmers, which can create unprecedented solutions. To display publicly all of its power and potential, IBM Watson attended the competition of the *Jeopardy!* Game in which it was able to show its ability [11], [12] and [13]. In fact, Watson was victorious and this is due to its ability in finding answers to questions in an unlimited domain. The Watson algorithms can mine both structured and unstructured data, extracting information and originating new knowledge, which results in the emergence of multiple answers that are weighted through a confidence estimation. The key feature required by the competition was mastering the language [14] because the game proposes clues full of irony, subtle meanings, and other complexities allowed by a malicious use of natural language [15]. For what concerns the underlying technology, we report that Watson owes its smartness to the DeepQA Project [16] a massively parallel probabilistic evidence-based architecture, which can also be adapted to different business applications and additional exploratory challenge problems including medicine, enterprise search, and gaming [17].

To conclude this short overview on cognitive computing capabilities, we highlight that the exploitation of cognitive computing services is education is a paramount field of application. Specifically, according to the smarter university model [18], cognitive computing-based applications and services should be adopted for administration and management, and learning activities as well. From the e-learning point of view, we report some experiences that witness how cognitive computing can be an accelerator for students' achievements, and a valuable support for the teachers. In particular:

(i) integrating cognitive computing services in software applications can strongly enhance students' performances in computer science classes;

(ii) studying cognitive computing behavior can lead to significant results in Artificial Intelligence (AI) related studies;

(iii) using a cognitive computing layer for digital interactions with students can enhance their performances and ease the teachers' job in managing both classes and learning materials.

The interested readers can find a comprehensive review of cognitive computing in education in reference [19]. Moreover, cognitive computing techniques allow managing big data and these two issues mixed together are influencing the learning process [20].

IV. The Video Analysis and Recommendation System

To the aim of testing the effectiveness of cognitive computing techniques in the application of semantics, the demonstrative prototype that we are going to describe was made exploiting videos from the course "Computer Science 101", which are made freely available on the website of the Stanford University through their Stanford openEdX e-learning environment. We have chosen this course due the authority of the provider and to the quality of the spoken English, since we want the automated systems to perform speech-to-text conversion. From a very technical point of view, the service is made through the pipelining of different services borrowed from the Watson catalog, according to the sequence shown by the block diagram depicted in Figure 1.

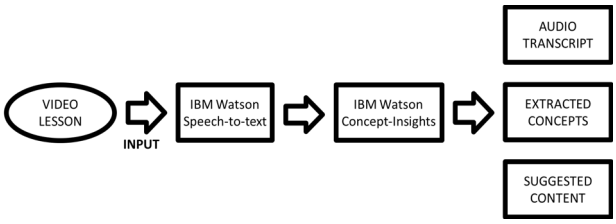


Figure 1: the sequence of operations

More precisely:

(i) a wide variety of videos is presented to the users;

(ii) the user is requested to select a specific video, which starts playing within the browser;

(iii) the video flows through a speech-to-text analysis, which generates its transcript;

(iv) the transcript outcome of the previous module flows through the *Concept Insights* module, which links the provided documents with a pre-existing graph of concepts based on Wikipedia;

(v) two types of links are identified: explicit links when a document directly mentions a concept, and implicit links which

connect your documents to relevant concepts that are not directly mentioned in them;

(vi) users can search for documents that are relevant to a concept or collection of concepts by exploring the explicit and implicit links;

(vii) the concepts identified in *(iv)* are ordered on the basis of the confidence level;

(viii) other related videos or web pages are suggested to the user.

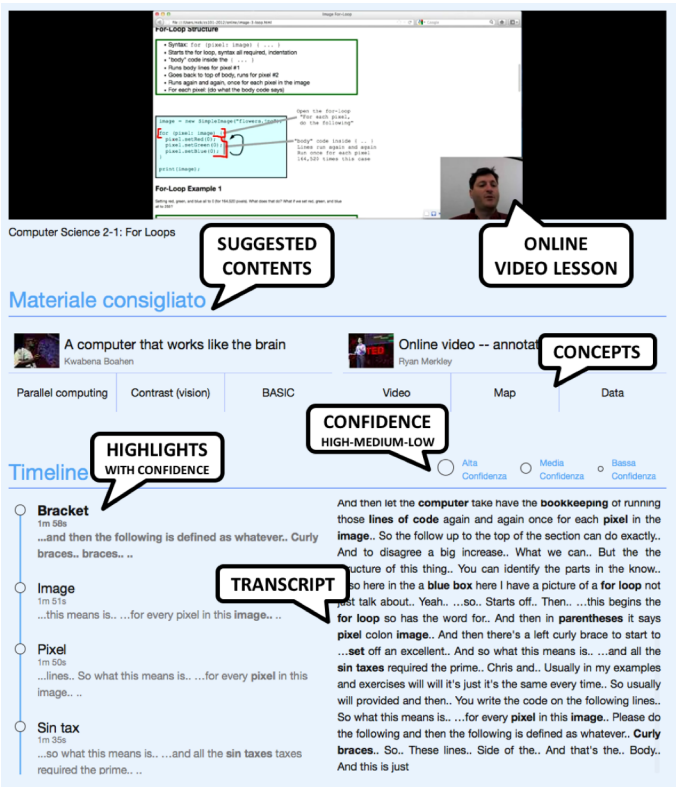


Figure 2: screenshot of the application while running, performing real-time transcript of text and the relevant semantic analysis, resulting in the suggested contents balloon

As a final result, data collected and originated by the analysis are presented in a single web page. This includes (see Fig. 2):

(i) the transcript of the video,

(ii) a timeline with a real-time overview of the keywords identified while the video plays,

(iii) a confidence estimation,

(iv) a list of the extracted concepts, and

(v) suggested related contents.

It is worthwhile noticing that these features can be useful also from an instructional designer point of view [21], which could make a video speaking about a topic, which will become the equivalent of the research keywords in standard search engines. In this way, basic concepts will be extracted from the video and these will become the table of contents. At the same time, the system will look for other related videos, which could enhance student engagement through extra credits (using gamification techniques [22] [23]) be adopted as learning materials in a class, be suggested as "further readings" [24] or re-arranged in a repository with semantic capabilities [25].

At your convenience, please note that the demo is available online at the URL https://mc-unige.eu-gb.mybluemix.net (best viewed with Mozilla Firefox).

V. Conclusions

Using videos in e-learning activity has become popular for both the availability of a large amount of video content and the increase of MOOCs projects carried on by universities, which push their researchers to support their teaching activity by publishing videos instead of presentations, books or documents. The use of videos itself is not an advancement in the e-learning methodology, since it was commonly used in the past, even if delivered in different forms, e.g., video-tapes. Yet, next generation network and services, advanced semantic capabilities and the availability of the cognitive computing services joined with the big data allow the empowerment of videos towards a new generation of e-learning frameworks able to overcome the traditional drawbacks of this medium.

Acknowledgments

We wish to thank Mr. Marco Placidi for his precious support in coding and setting up the demo application.

References

[1] Alexa, "Top 500 sites on the web," Annual Report available at the URL: http://www.alexa.com/topsites, last visited, September, 2015.

[2] E. Garcia-Barriocanal, M.A. Sicilia, S. Sanchez-Alonso, M. Lytras, "Semantic annotation of video fragments as learning objects: a case study with YouTube videos and the Gene Ontology." Interactive Learning Environments, vol.19, no. 1, pp. 25–44, 2011.

[3] J.K. Tang, H. Xie, H., T.L. Wong, "A big data framework for early identification of dropout students in MOOC," in Technology in Education. Technology-Mediated Proactive Learning, Springer Berlin Heidelberg pp. 127–132, 2015.

[4] K.S. Hone and G.R. El Said, "Exploring the factors affecting MOOC retention: A survey study," Computers & Education, vol. 98, pp. 157–168, 2016.

[5] A. Elbadrawy, A. Polyzou, A. Ren, M. Sweeney, G. Karypis, H. Rangwala, "Predicting Student Performance Using Personalized Analytics," Computer, vol. 49, no. 4, pp. 61–69, 2016.

[6] K. Kobylinski, J. Bennett, N. Seto, G. Lo, F. Tucci, "Enterprise application development in the cloud with IBM Bluemix," in Proceedings of the 24th Annual International Conference on Computer Science and Software Engineering (CASCON '14), pp. 276–279, 2014.

[7] J.E. Kelly, S. Hamm, "Smart machines: IBM's Watson and the era of cognitive computing." Columbia Business School Publishing, 2013.

[8] M. Coccoli, P. Maresca, L. Stanganelli, A. Guercio, "An experience of collaboration using a PaaS for the smarter university model," Journal of Visual Languages and Computing, vol. 31, no. 1, pp. 275–282, 2015.

[9] G. Booch, "The soul of a new Watson, IEEE Software," vol. 28, no. 4, pp. 9–10, 2011.

[10] S. Sudarsan, "Evolving to a new computing era: cognitive computing with Watson," Journal of computing sciences in colleges archive, vol. 29, no. 4, pp. 4–4, 2014.

[11] A.M. Gliozzo, "IBM's Watson Jeopardy! computer shuts down humans in final game," IEEE Spectrum, vol.17, 2011.

[12] E.W. Brown, "Watson: the Jeopardy! challenge and beyond," in Proceedings of the 35th International ACM SIGIR conference on research and development in information retrieval, 2012.

[13] E.W. Brown, "The Jeopardy! challenge and beyond," in Proceeding of the 12th IEEE International conference on Cognitive Informatics & Cognitive Computing (ICCI*CC), 2013.

[14] M.C. McCord, J.W. Murdock, B.K. Boguraev, "Deep parsing in Watson," IBM Journal of research and development, vol. 56, no. 3.4, 3:1, pp. 3:15, 2012.

[15] A. Lally, J.M. Prager, M.C. McCord, B.K. Boguraev, S. Pathwardhan, J. Fan, J. Fodor, J. Chu-Carrol, "Question analysis: how Watson reads a clue," IBM Journal of research and development, vol. 56, no. 3.4, pp. 2:1, 2:14, 2012.

[16] D. Ferrucci, A. Levas, S. Bagchi, G. Gondek, E.T. Mueller, "Watson: beyond Jeopardy!," Artificial Intelligence, 199–200, pp. 93-105, 2012.

[17] E.A. Epstein, M.I. Schor, B.S. Lyer, A. Lally, E.W. Brown, J. Cwiklik, "Making Watson fast," IBM Journal of Research and Development, vol. 56, no. 3.4, pp. 15:1-15:12, 2012.

[18] M. Coccoli, A. Guercio, P. Maresca, L. Stanganelli, "Smarter universities. A vision for the fast changing digital era," Journal of Visual Languages and Computing, vol.25, no.6, pp. 1003–1011, 2014.

[19] M. Coccoli, P.Maresca, L. Stanganelli, "Cognitive computing in education," Journal of e- Learning and Knowledge Society, vol.12, no.2, pp. 55–69, 2016.

[20] M. Coccoli, P. Maresca, L. Stanganelli, "The role of big data and cognitive computing in the learning process," Journal of Visual Languages and Computing, In press.

[21] M. Coccoli, S. Iacono, G. Vercelli, "Applying gamification techniques to enhance effectiveness of video-lessons", Journal of e-Learning and Knowledge Society, vol. 11 no. 3, pp. 73–84, 2015.

[22] C. Perryer, N.A. Celestine, B. Scott-Ladd, C. Leighton, "Enhancing workplace motivation through gamification: Transferrable lessons from pedagogy," The International Journal of Management Education, vol. 14, no. 3, pp. 327–335, 2016.

[23] G. Adorni, S. Battigelli, D. Brondo, N. Capuano, M. Coccoli, S. Miranda, F. Orciuoli, L. Stanganelli, A.M. Sugliano, G. Vivanet, "CADDIE and IWT: two different ontology-based approaches to Anytime, Anywhere and Anybody Learning," Journal of e-Learning and Knowledge Society, vol. 6, no. 2, pp. 53–66, 2010.

[24] M. Coccoli, G. Vercelli, G. Vivanet, "Semantic Wiki: a collaborative tool for instructional content design," Journal of e-Learning and Knowledge Society, vol. 8, no. 2, pp. 113–122, 2012.

Teaching Computer Programming Through Hands-on Labs on Cognitive Computing

Mauro Coccoli
DIBRIS
University of Genoa
Genoa, Italy
mauro.coccoli@unige.it

Paolo Maresca
DIETI
Federico II University
Naples, Italy
paolo.maresca@unina.it

Lidia Stanganelli
Software Engineer
Naples, Italy
ldistn@gmail.com

***Abstract*—In this work we report more recent results of a long-lasting educational project that we have been carrying on for several years and is evolving continuously. The objective of the mentioned project is making students work on the production of small, yet full featured, software prototypes with a collaborative approach, in a dedicated development environment with suited tools and facilities. At the same time, we seek to lay the foundations to build a pragmatic model to teach cognitive computing programming. We conducted such experience in a programming course at the University of Naples "Federico II" exploiting a software development environment based on the Platform as a Service (PaaS). This made possible to perform a set of cooperative learning activities that we used to demonstrate practically some theoretical concepts, also stressing the use of cognitive computing tools, which introduce a new way of thinking the software design process. From its inception, this educational project has involved a relevant number of students, which has been growing continuously, year after year. At our first attempts, students were assigned a specific activity, which was expected to be concluded within one year. Then, as new classes arrived, we have seen that it was possible to start new activities starting from the precedent achievements, thus requiring further evolutions of the available prototypes. This allowed creating more complex and complete projects, as new tools and services were made available, carrying new opportunities. In the present release, such evolutionary path has led to using the IBM Bluemix platform with its wide range of components, including Watson that is devoted to cognitive computing. This work goes in the direction of developing the smart university model, by using innovative and intelligent services to help raising a new generation of software engineers but also to promote and disseminate a new way for designing and building innovative applications.**

Keywords: cloud; cognitive computing; collaborative systems; smart university; technology enhanced learning; computer supported collaborative learning.

I. INTRODUCTION

As already stated in previous research, collaboration is a learning strategy that can be applied in educational activities specifically designed to bring learners to a given learning outcome throughout collaborative tasks. By adopting collaborative methodologies, teachers can empower the acquisition of skills, competence and knowledge, thus reinforce the whole learning process, also enhancing mental capabilities of individuals, driving them towards acquiring a cooperative attitude and sharing resources with peers. Especially, collaboration is a paramount aspect in computer-assisted instruction, resulting in the implementation of the Computer-Supported Collaborative Learning (CSCL) paradigm [1]. This also reflects in the architecture of widely adopted e-learning platforms, that include functionalities specifically developed for allowing users to collaborate and/or to cooperate, to the aim of reaching common objectives.

Furthermore, in specific disciplines, such as, e.g., in software engineering classes, collaboration is also one of the learning objectives. In fact, software development is done by heterogeneous people arranged in teams including designers, analysts, programmers, testers, user experience designers, graphic designers, user interface designers, and many others, also depending on the dimensions of the projects. All these people must share common objectives, have a common and clear vision of the project, and be able to communicate each other seamlessly. Thus, a suited working environment is required, which includes some non-standard functionalities that may be not present in common Learning Management Systems (LMS) since they usually serve for content delivery and accounting of users. Moreover, we notice that making concrete experimentations in laboratories equipped with suited facilities, is one of the core components of education along the fan of the scientific disciplines [2]. This is due to the fact that experimenting enables students to learn by doing through simulations and hands-on experiences, which are at the basis of many practical educational activities. Specifically, considering the case of teaching software engineering, the laboratory equipment is not made of machineries, yet it is all about computers, networks, operating systems, IDEs (Integrated Development Environments) and software programs, which can even be conveniently virtualized [3].

Within this context, we have been developing, appositely for computer programming classes, a dedicated working environment, which enables students, alone or in small groups, working on small and simple tasks within more complex software engineering projects, covering the whole lifecycle of software development across analysis, design, and

DOI reference number: 10.18293/DMS2016-023

development. Furthermore, this architecture was also designed to allow for subsequent upgrades by other students in future years. Such a working environment was designed with reusability in mind and is continuously evolving, as new solutions emerge, which can be profitably exploited by both students and teachers. This reflects positively on learning methodology, fostering an innovative model for computer programming classes, but also results in dramatic changes in universities from an infrastructure and organizational point of view [4]. In fact, one of the main issues in making available a full featured flexible, sharable and reusable working environment is the fact that it clashes with the static structure of existing software engineering laboratories. Traditionally, universities have invested much money, and human resources as well, in setting up and maintaining large and costly computer rooms with fast Internet connections for their students' satisfaction. However, it is evident that this model has many drawbacks such as, e.g.,

(i) the need to be reconfigured according to the specific requirements that may come from different courses;

(ii) the need of keeping updated, consistent and protected the software installed as security patches or new versions are released;

(iii) many security and safety threats;

(iv) the need of specific and costly hardware supporting huge loads, even if concentrated in short periods, i.e., the few "lab-hours" scheduled during a semester;

(v) the rapid obsolescence of hardware.

All of these issues can be summarized in just one point: high costs. Moreover, in recent years, other inefficacy reasons emerged, due to technology progresses and changes in the society and in people's lifestyle, habits and attitudes. In fact, modern construction techniques guarantee ever-growing computing power and progressive miniaturization rates, which result in affordable and capable devices. In addition, the fast and continuous innovation in network infrastructures and networking solutions [5] results in the straight connection between people and objects and in the rapid development of Internet of Things (IoT) applications and solutions, also bringing innovation to technology-enhanced learning [6]. Furthermore, the wide availability of applications for sensor-rich modern devices pushes to using methods and techniques strongly based on the interaction with users, which happens mainly through natural interfaces.

Accordingly, in this paper we focus our attention on a specific issue, that is, the exploitation of the Platform-as-a-Service (PaaS) paradigm in the software production [7]. The adoption of PaaS allows using remote virtual machines in place of local hardware and software, thus avoiding time-consuming and expensive installations as well as annoying maintenance tasks [8]. Then, we focus on cognitive computing that, together with big data, is going to assume a fundamental role in the learning process [9] and is being rapidly adopted in a variety of education settings [10]. Moreover, cognitive computing is a very challenging item in the software development scenario since it allows developing applications that manage big data, enables performing analytics on data in order to make complex decisions, and fosters interaction between machines and persons through natural language, both written and spoken. Cognitive computing will be increasingly linked both to IoT and analytics because of the large amount of data collected and the subsequent decision that can be taken once these data will be analyzed [11]. Especially, if we combine the pattern matching with the complex verbal communication, training students in our labs on these issues is necessary and important for the future applications and the formation of software engineers.

The remainder of the paper is organized as follows. First, in Section II, we present a summary of past experiences, which have allowed reaching the actual situation. Then, in Section III we introduce general issues about cognitive computing. In Section IV we detail on the experiments carried on with the IBM Watson cognitive computing facilities into several projects. Finally, in Section V, summary considerations conclude the paper.

II. A Summary of Past Experiences

In this section, we report on the previous experiences that improved collaboration between students and drove the project to the current state. This evolution path must not only be seen as a recasting of the instruments used, which also served to enhance collaboration, but should also be regarded as the maturation of the way in which we obtained a strong evolution in the control management and implementation processes of collaboration between people from diverse groups and countries.

A. Enforcing Team Cooperation

In March 2010, a first project started from the collaboration between the University of Naples Federico II, IBM Italy, the Eclipse Italian community, and other Italian universities. The aim of the project was the collaborative creation of good-quality software products within university courses, to improve students' learning and to enhance the collaboration capabilities between members of the development teams. This resulted in a new way of teaching and bridged the gap between companies and universities, with the objective of preparing students to face at the best the labor market. The project was called ETC (Enforcing Team Cooperation) and it was attended by the following universities: Naples Federico II, Bergamo, Milano Bicocca, Genoa, Bari, Bologna. The project involved one hundred students and used the IBM Rational's Jazz vision (see references [12] and [13] for more detailed descriptions).

B. On the Road to Eclipse

Since the ETC methodology was successful, it was applied again one year later, 2011, in a larger project called OTRE (On the Road to Eclipse) [14] for developing a set of collaborative tools, to be used within the Eclipse community. The final results were collected, published and shared, in a workshop in which the teams that had collaborated remotely finally met and were able to compare their work with other teams involved in the same project. Something like a peer-to-peer conference, that wanted to be much more than a student contest.

C. *Trans-Continental Collaboration*

Encouraged by the good results achieved, the next year, 2012, we faced an ambitious trial centered on cooperation of student groups belonging to both Italian and non-Italian universities [15]. The project involved two classes: a programming class and a computer science class. The former from the Federico II University of Naples (Italy) and the latter from the Ohio State University at Stark (USA). The goal was to form the heterogeneous classes on a platform (Rational-Jazz) and set up a 24 hours workshop that students could use at any time. As a consequence, a common training on common educational objectives could be achieved by standardizing the programs and using the same language.

D. *Geographically Distributed Teams*

The outcome of this trial was very satisfactory for both teachers and students, hence it was extended and repeated the next year, 2013. New projects were set up with mixed teams that represent portions of both Italian and American students. This made it possible to tackle problems in training places such as, e.g., difficulty in speaking, different study habits, organizational habits, etc. [16]. We are proud to mention that this project has received a special mention by IBM, as the best practice in collaboration of the year[1].

III. Cognitive Computing and Watson

The above-cited experiences contributed to the definition of the model of smarter university given in [4], which depicts the progresses that universities are expected to do (or are already doing) for providing their students with smart services. These are implemented in remote laboratories so that they will be available everywhere, for all time and without the need of reservation. This concept finds its practical realization in the newly acquired possibility of using a set of cloud applications in a PaaS environment, i.e., the IBM Bluemix platform [7]. With this step ahead, students are enabled to cooperate, develop, maintain and manage any project with any team at any time without the need to move to any physical laboratory. The philosophy of virtual all without any installation in any laboratory or PC and the ability to use hardware resources and software for free (owing to the IBM Academic Initiative) make this a very powerful service. Finally, we highlight that all of these premises are changing the programming paradigms and are driving applications to demand more cognitive services.

Cognitive computing can be regarded as the computer simulation of the human reasoning processes. To give a more detailed definition, we must make some considerations on the concept of programming. More precisely, we recall that the era of programmable computing starts with the dawn of the digital computer around 1940. The big change introduced is that now one has general purpose computing systems that are programmable and that can be reprogrammed to perform different tasks, according to the needs of the moment. But ultimately, they have to be programmed and are still somewhat constrained in the way they interact with humans. This is the era that we are currently in. However, we see emerging over the last few years what we call cognitive computing, as the result of summing a number of factors. In fact, there are at least four issues that affect the cognitive computing:

(i) big data,

(ii) expanding human cognitive boundaries,

(iii) improving communication, and

(iv) analytics.

Each of these is briefly described in the following.

A. *Big Data*

The first factor is the emergence of big data, originated by ever-growing number of services and applications that we daily use over the World Wide Web and the Internet, exchanging information with both humans and computers, and creating new knowledge. Apparently, we are not prepared to handle this huge amount of data and this results in using them rarely, wasting big opportunities and valuable resources.

B. *Expanding Human Cognitive Boundaries*

The second factor is due to the need to amplify the human cognitive boundaries. It is unquestionable that our ability to reason and think deeply, as well as to solve complex problems, is really quite impressive. However, our ability to read, analyze and process huge volumes of data is really quite poor.

C. *Improving Communication*

The third factor is due to the need to easily communicate to both machines and persons through natural language. So, instead of writing and editing programs that continuously interact with machines we can just talk and educate them about things to do. At the same way, machines can talk to other machines or humans in a dynamic always-connected IoT scenario.

D. *Analytics*

Finally, the fourth factor, which can be seen as a substrate, is analytics. This factor is more properly the key element of cognitive computing applications and is represented by the intense use of machine learning [17].

Figure 1. Some of the Watson boilerplates

These four enablers push the development of cognitive computing and relevant applications. Probably, there will be a transitional, maybe long, period in which traditional programming and cognitive programming will coexist and, sometimes, will overlap. According to this outlook, one can

[1] IBM Academic Award http://ibm.co/1Og2YNt

understand why the IBM Bluemix platform and Watson co-exist. In fact, they are the way to share the two above-described programming paradigms. Figure 1 shows the Watson services available within the Bluemix catalog.

In more detail, Watson is a full-featured cognitive computing system for the research and development of cognitive systems and services [18], [19], which can easily interoperate with other applications and legacy systems as well. Owing to the IBM Academic Initiative, universities can freely experiment Watson through the ever-growing variety of services and APIs (Application Programming Interfaces), made available within the IBM Bluemix platform [20], which represent an important and promising field of application for the exploitation of cognitive computing services is education [21]. Some of the many available Watson boilerplates will be used within the programming course for the creation of *cognos* for smartphones and PCs as well. For example, the most used are: *Dialog*, *Text to Speech*, *Speech to Text*, *Visual Recognition*, *Language Translation*, and *Personality Insights*.

IV. Outcomes of Hands-on Lab Activity

During the course of *Programming I*, in the academic year 2015/2016, the Federico II University of Naples launched the second trial with IBM Bluemix. Specifically, this year, the students carried on 14 different projects. Among these, only a small number are finished running while others are still under development, but are nearing completion. The students involved are 150, working in 50 groups for crafting applications for both smartphones/tablets and PCs. As in previous years, the students were divided into teams and each of these was assigned a part of a bigger project.

To better clarify what we mean with "students' project", we report the ones completed, which are currently in testing phase. The projects are, namely: *(i) Bag of ideas, (ii) Sbulloniamoci, (iii) Cancer registry, (iv) Restart Campania, (v) Arduino car.*

A. Bag of Ideas

This application resulted in the creation of a Social Network Site (SNS) for searching opportunities and sharing ideas among creative people who want to develop innovative services with some interesting features. In other words, the SNS aims to provide potential customers with the possibility of requesting assistance and expertise. More specifically, it works as follows: if you have an innovative idea but not the ability to achieve it, then you can post it on the site and ask for help (see Fig. 2). The SNS administrators make available skills and means to realize the idea. From a logical point of view, the core of the system is the analytics service that collects data, performs analyses and, finally, provides an answer on the feasibility of the idea, keeping into account the technological drivers, the territorial requirements, the existence of similar ideas, etc., applying suited analytics algorithms. Moreover, it also provides an application allowing people to manage the reputation of the participants through a peer rating system operated within the community of registered users. The users' reputation is measured ranging from 0 to 10.

B. Sbulloniamoci

The main purpose of this web application is mining contents of the most popular SNSs (e.g., Facebook and Twitter, also exploiting their known privacy flaws [22], [23]) investigating possible threatening messages, in order to recognize the profile of a bully and report it to the community. This concept is already implemented in the popular application ReThink, which keeps trace of outgoing messages. Whenever a message contains threats, the application displays the message «Are you sure you want to do that?» to discourage unfair behaviors.

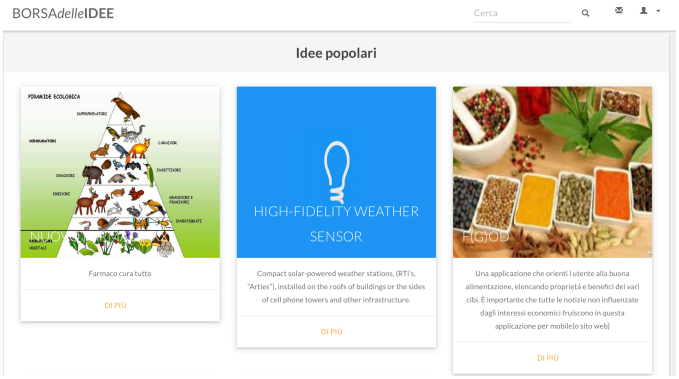


Figure 2. Screenshot taken from Bag of ideas

However, *Sbulloniamoci* (Fig. 3) wants to be more incisive and the result of writing malicious messages is not just a warning but the exposition of the evil person to the scorn.

From a technical point of view, the application performs two steps. The former is the interaction with the user, achieved by means of the *Dialog* component from Watson. The latter, which is the application's core functionality, is the elaboration of the text that may result in the unmasking of the "bully". This phase is performed by *Personality* from Watson. It allows analyzing a text to the aim of extrapolating the writer's psychological profile, showing its main characteristics, as well as what actions he may or may not accomplish. Note that, in order to achieve significant results, *Personality* needs sets of 3500 to 6000 words. Using few words will reflect in getting less reliable profiles. Before using *Personality*, it can be very useful text through the facility *Concept Expansion* and *Concept Insights* also available within the Watson libraries.

Figure 3. Screenshot taken from Resize the bully – "Sbulloniamoci"

In fact, *Concept Expansion* helps clarifying youth languages or slang in more clear terms (for example, it makes a connection between the words *Big Apple* and *New York* that will be considered synonymous) and such a pre-processing can be very useful to better understand the users' intentions. Moreover, *Concept Insights* allows linking given words to concepts; thus, it can detect any threats, insults, or aggressive language that may not be detected by merely scanning the text.

C. *Cancer Registry.*

The main aim of the *Cancer Registry* project is to build a website and an app to show the data for the more occurring tumors in a given geographical area. The motivation behind this project is finding a relationship between such serious diseases and pollution (air, water and waste management in general), which is supposed to increase cancer presence in specific areas of the country. This huge work uses data from disparate sources, i.e., the National Cancer Registry of Italy, environment sensors monitoring air and water pollution, geographical data from the Google maps service. The development was assigned to 4 separate groups as follows:

(i) database;

(ii) data analysis;

(iii) application development (currently available only for devices running the Android OS, it will be built for iOS soon);

(iv) website development.

It is worthwhile noticing that, in this case, it was hard identifying a specific boiler plate tailored to our needs. We are considering the possibility to develop an ad-hoc component based on Watson.

Owing to the *Analytics* functionalities provided, it is possible analyzing data from different sources. In the chart (Fig. 4), the age of patients hospitalized for tumors in the 2003-2013 decade is depicted. From this chart we can find the age with more incidence of tumors in the selected geographical area.

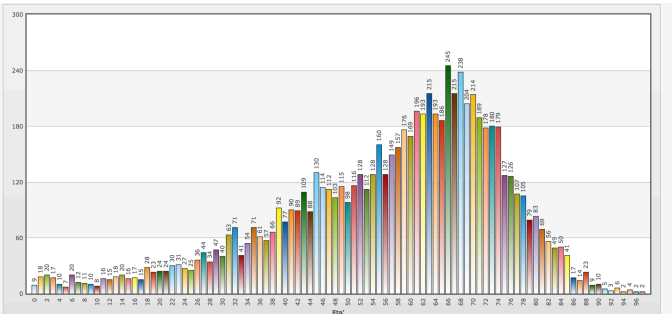


Figure 4. Patient age hospitalized for tumors in the 2003-2013 decade

D. *Restart Campania.*

The *Restart Campania* project is the development of an application that allows promoting tourism in the Italian region Campania. One of its core functionalities is visual recognition, which enables users, i.e., tourists and citizens, to adopt the places and monuments they take picture of, by making a donation or by taking part to the activities of a community devoted to promote a specific point of interest, to attract more and more visitors, thus enhancing culture, tourism and commerce. The application implementing this service includes an app for smartphone and/or tablet and a website. As users take photos of specific monuments, the system creates a map with the relevant markers. Specifically, images are recognized through the *Visual Recognition* tool, which looks for a match in the huge database of images it relies on. Once the place or monument in the picture is identified, the app performs a Google search to provide the user with more information. Moreover, the monument is highlighted on a Google map (see Fig. 5) where corresponding markers are linked to videos uploaded by users. (see Fig. 6).

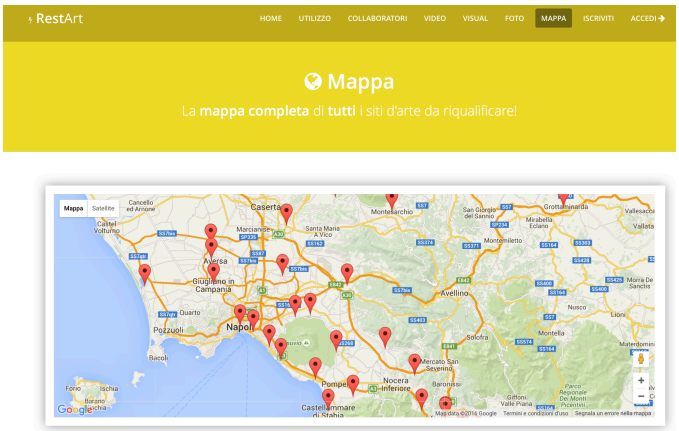


Figure 5. Map of adoptable monuments

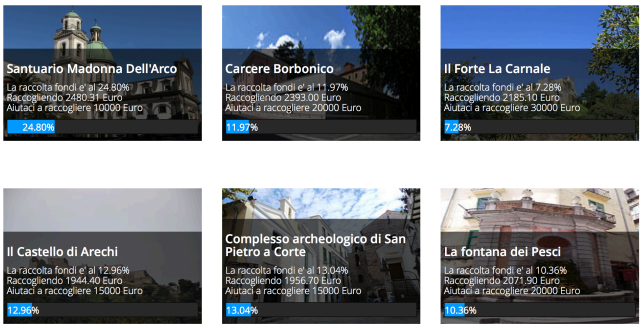


Figure 6. Some of the monuments and details on the relevant donations

E. *Arduino Car*

The *Arduino Car* project received the special mention by IBM during the last Eclipse-IT workshop, "Head in the Clouds", held in Rome on October, 2015. In more detail, Arduino Car is the design and development of a real, working, toy-machine and its remote control system, running in a smartphone app. It was entirely developed within the IBM Bluemix PaaS and exploiting Watson functionalities, including specific IoT facilities. Figure 7 shows the prototype at its current state of development.

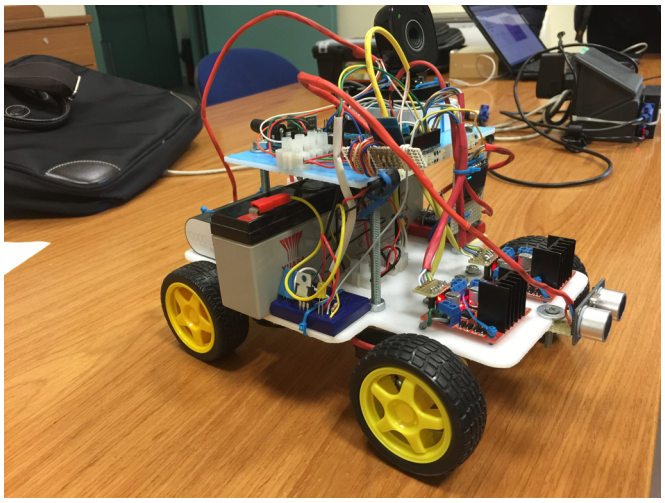


Figure 7. Arduino Car setup

From an engineering point of view, the project is interesting since it involves three different activities: the human-machine interaction, the realization of the motor and the control of all the on-board devices, including traction control and a variety of sensors. As regards the first aspect, the students have developed a simple communication protocol, which allows fast communication (also because of the slowness of the communication network) and they enabled voice commands using the speech-to-text functionality. For mechanical aspects, it is worthwhile noticing that the electric motor has been designed and assembled so that traction can be on two or four wheels by engaging/disengaging one of the axis, i.e., front and/or rear. The last aspect, the vehicle control system, required a huge software development activity, for the controller is in charge of managing both the electronic equipment on board (e.g., light switching, distance control sensors, temperature sensors, etc.) and the navigation on the road as well as collision avoidance.

In addition, a remote dashboard was designed to control the overall performances and the functionality of the car via smartphone, PC or tablet. Such dashboard also displays the video stream of a camera installed on the front of the vehicle. Another camera on the back allows controlling the vehicle in parking with the help of sensors. In more detail, it carries 4 different mini-cameras, heat sensors and distance sensors. The software developed gathers data, analyzes them and provides support to the decisions app that monitors the car's performance. The working group that developed the project has had to deal with the hardware, the on-board software, the communication network, also developing a communication protocol between Watson and the control system.

Future developments of this project will be: enhancing the equipment (easing end empowering), enhancing aerodynamics, and improving vehicle coordination with the surrounding environment and with other vehicles. Due to its interdisciplinarity, the project also involved students enrolled in the course of studies in mechanical engineering for engine design and on-board electrical systems to control the 2-wheels drive traction system.

V. Conclusions

In conclusion, we observe that teaching computer programming can be performed through a variety of strategies and based on different languages or IDEs, according to the teachers' preferences and to the available facilities in terms of both software licenses and hardware. However, due the practical nature of such a discipline, we noticed that better results are achieved when students have to deal with realistic problems and realize real working applications. This cannot be attained starting programs from the scratch but requires the ability of using collaborative methodologies and of reusing libraries already available, with an incremental approach to the construction of full-featured software prototypes. This is also the requirement of the labor market, which needs flexible professionals with problem solving capabilities and social attitudes. Using the ecosystem introduced within the paper, i.e., the IBM Bluemix PaaS and the relevant services, in hands-on labs of the programming course has guaranteed a high quality of the final artifacts that were used for evaluation by the teachers. Moreover, through the PaaS, we can exploit advanced functionalities based on a very complex cognitive computing system. Such services, perform high-level operations and require high investments in terms of hardware computational capabilities and research and development as well, that could not be faced in a university class. Since the platform makes them available for free, this can empower all the applications with high level functionalities at no cost, also enhancing the satisfaction level of the students and their performances. Consequently, this reflects on a better performance from the teacher side too.

Acknowledgments

The authors wish to thank IBM and the Italian Eclipse Community for their friendship and their kind collaboration.

References

[1] B.G. Silverman, "Computer Supported Collaborative Learning," Computers Education, vol. 25, no. 3, pp.81-91, 1995.

[2] L. Caviglione, M. Coccoli, E. Punta, "Education and training in grid-enabled laboratories and complex systems," in Remote Instrumentation for eScience and Related Aspects, pp. 145-157, Springer, 2012.

[3] R. Uhlig, G. Neiger, D. Rodgers, A.L. Santoni, F.C.M. Martins, A.V. Anderson, S.M. Bennett, A.Kagi, F.H. Leung, L. Smith, "Intel virtualization technology," Computer, vol. 38, no. 5, pp. 48-56, 2005.

[4] M. Coccoli, P. Maresca, L. Stanganelli, A. Guercio, "Smarter universities: a vision for the fast changing digital era," Journal of Visual Languages and Computing ,vol. 25, no.6, pp. 1003-1011, 2014.

[5] L. Caviglione, M. Coccoli, V. Gianuzzi, "Opportunities, integration and issues of applying new technologies over e-learning platforms," in Proceedings of the 3rd International Conference on Next Generation Networks and Services, NGNS'2011; Hammamet; Tunisia; Dec. 18-20, 2011, pp. 12-17.

[6] G. Adorni, M. Coccoli, I. Torre, "Semantic web and Internet of Things supporting enhanced learning," Journal of e-Learning and Knowledge Society, vol. 8, no. 2, pp. 23–32, 2012.

[7] M. Coccoli, P. Maresca, L. Stanganelli, A. Guercio, "An experience of collaboration using a PaaS for the smarter university model," Journal of Visual Languages and Computing, vol. 31, no. 1, pp. 275-282, 2015.

[8] G. Lawton, "Developing software online with Platform-as-a-Service technology," Computer, vol. 41, no. 6, pp. 13–15, 2008.

[9] M. Coccoli, P. Maresca, L. Stanganelli, "The role of big data and cognitive computing in the learning process," Journal of Visual Languages and Computing, In press.

[10] M. Coccoli, P. Maresca, L. Stanganelli, "Cognitive computing in education," Journal of e-Learning and Knowledge Society, vol. 12, no. 2, pp. 55–69, 2016.

[11] J.E. Kelly and S. Hamm, "Smart machines: IBM's Watson and the era of cognitive computing,". Columbia Business School Publishing, 2013.

[12] M. Coccoli, P. Maresca, L. Stanganelli, "Enforcing team cooperation: an example of computer supported collaborative learning in software engineering," Proceedings of the 16th Interantaional Conference on Distributed Multimedia Systems, Workshop on Distance Education Technologies, 2010, pp. 189-192.

[13] M. Coccoli, P. Maresca, L. Stanganelli, "Computer Supported Collaborative Learning in software engineering: (ETC - Enforcing Team cooperation)," Proceedings of the 2011 IEEE Global Engineering Education Conference, EDUCON 2011, Amman, Jordan; April 4-6 2011, pp. 990-995.

[14] P. Maresca, L. Stanganelli, M. Coccoli, "Managing a software project leveraging students' cooperation: on the road to Eclipse (OTRE) experience," in Proceedings of 12th International Conference on Product focused software development and process improvement: PKMT-Project and knowledge management trends, Torre Canne (Br), Italy, June 20-22, 2011, pp. 96-100.

[15] P. Maresca, A. Guercio, L. Stanganelli, "Building wider team cooperation projects from lessons learned in open communities of practice," in Proceedings of the 18th International Conference on Distributed Multimedia Systems (DMS2012) – Distance Education Workshop, Miami Beach, USA, August 9-11, 2012, pp. 144-149.

[16] P. Maresca, A. Guercio, L. Stanganelli, "Modeling multiple common learning goals in an ETCplus educational project", in Proceedings of the 19th International Conference on Distributed Multimedia Systems (DMS2013) – Distance Education Workshop, Brighton, United Kingdom, August 8-10, 2013, pp. 122-127.

[17] S. Earley, "Executive roundtable series: machine learning and cognitive computing," IT Professional, vol. 17, no. 4, pp. 56-60, 2015.

[18] G. Booch, "The soul of a new Watson," IEEE Software, vol. 28, no. 4, pp. 9-10, 2011.

[19] S. Sudarsan, "Evolving to a new computing era: cognitive computing with Watson," Journal of computing sciences in colleges archive, vol. 29, no. 4, pp. 4-4, 2014.

[20] K. Kobylinski, J. Bennett, N. Seto,G. Lo, F. Tucci, "Enterprise application development in the cloud with IBM Bluemix," in Proceedings of the 24th Annual International Conference on Computer Science and Software Engineering (CASCON '14), Markham, Canada, 2014, pp. 276-279.

[21] M. Coccoli, P. Maresca, L. Stanganelli, "Cognitive computing in Education," Journal of e-Learning and Knowledge Society, vol. 12, no. 2, pp. 55-69, 2016.

[22] L. Caviglione and M. Coccoli, "Privacy problems with Web 2.0," Computer Fraud and Security, vol. 2011, no. 10, pp. 16-19, 2011.

[23] L. Caviglione, M. Coccoli, A. Merlo, "A taxonomy-based model of security and privacy in online social networks," International Journal of Computational Science and Engineering, vol. 9, no. 4, 2014, pp. 325-338, 2014.

A Context-Aware Approach for a Collaborative, Pervasive and Adaptive Digital Storytelling

M. Casillo
DII
Università di Napoli "Federico II"
Napoli, Italy
mario.casillo@unina.it

F. Colace, M. De Santo
DIIn
Università degli Studi di Salerno
Fisciano (SA), Italy
{fcolace, desanto}@unisa.it

S. Lemma, M. Lombardi
SIMASLab
Università degli Studi di Salerno
Fisciano (SA), Italy
{slemma, malombardi}@unisa.it

Abstract

The paper introduces an approach for the realization of a dynamic and contextual storytelling engine that can allow the dynamic supply of narrative contents, not necessarily predetermined and pertinent to the needs and the dynamic behaviors of the users.
In particular, the system adopts an adaptive and social approach, using a contextual model in order to realize a dynamic digital storytelling approach that makes the visit experience more appealing and immersive. A case study and some experimental results are presented and discussed.

1. Introduction

"By dint of telling his stories, a man becomes the stories. They continue to live after him, and so he becomes immortal" (from "Big Fish" film).

The mystery contained in each "word" is still one of the most debated issues in the entire history of humanity. From this mystery comes the narration and its ability to persuade, to convert, to enchant [3, 24].

With the term *'storytelling'*, we mean a learning methodology that consists in 'telling' a story to catch the attention of an audience, convey the message that the story wants to tell and stimulate a specific desire in the readers, transmitting events in words, images and sounds [1].

The objectives of storytelling are several and vary according to the context of application: in fact, the disciplinary areas are many and go from the literacy in childhood, to the scholastic education and the museum didactics, passing for other uses such as the company management and the psychoanalysis [26].

If today we can think of using a didactic methodology as the storytelling, it is mainly thanks to two factors: from the one hand, a new way to imagine the education seen from a constructivist perspective, on the other hand the diffusion in the latest decades of new technologies [15].

In this context, we talk about 'digital storytelling', of which there are several definitions, but all turn around the 'idea to combine the art of telling stories with the variety of digital multimedia services, such as images, audios and videos' [23]: mixture that allows expressing and narrating in a vivid way experiences, situations and considerations.

Digital storytelling exists in many forms and encompasses multiple fields; in particular, there are the following typologies: linear, non-linear, adaptive, social/collaborative, mobile and game.

Linear storytelling is a traditional narrative form where the author chooses a predefined plot with a fixed sequence of events that the protagonist cannot modify in any way. The term 'linear' indicates that any part of the content is seen and/or listened to in the same order all the times it is enjoyed, confirming the fact that the reader is pushed to follow the narrative path without the possibility to influence the plot and/or its end [28].

Non-linear storytelling represents a 'not linear' narration form, that is any corpus of contents structured in a way that all the possible ways to cover are multiple and variable [19].

The main idea is that the advancement of the plot is not fixed a priori, but the development of the story and its realization depend on the interaction and the choices of the audience, who from simple reader or spectator becomes a 'co-author'.

Instead, the adaptive approach is based on an interactive narration that allows intervening and interacting with the process of construction of the story adapting it to the alterations caused by the intervention of the user, with the purpose to keep a certain narrative coherence [12].

The use of media and social media for the sharing of contents and narrative artefacts has considerably transformed even the way in which the stories can be described and told.

'Collaborative' and 'social' storytelling are forms of collective construction of a story where the shared narration becomes a socially interactive means that supports and foments the continuous co-participation of the users [9].

If the use of technology has changed the way to tell the stories, the use of new mobile tools has put the problem on the overlapping of the narratological and the spatial parts. Consequently, the researchers and the developers have been pushed to define new approaches able to integrate the two dimensions in order to make the cultural experience as much as possible immediate and at the same time ubiquitous.

Many projects use the mobile component to realize guided narrative itineraries contextualized to the environment, with an instructive, informative or sometimes simply recreational feature, with active users who move

DOI reference number: 10.18293/DMS2016-033

inside physical places operating on technological devices [14].

Also the support of geolocation (positioning and localization through integrated GPS) has further improved the use of these mobile tools, allowing a strong reinforcement of the experience offered to the visitor. In fact, it is very current the theme of the creation of geolocated interactive narrations aware of the context, that is, of stories whose contents, organized in a narrative sequence, are strongly anchored to the physical places and to other significant parameters of the environmental context [6, 7, 20].

Therefore, it is understandable how it is essential to give the citizen adaptive services and contents able to transmit right information in the right contexts [33]. A fundamental contribution to this scenery has come from the increasing diffusion of mobile devices that concretely have created pervasive and context-aware computing approaches [21, 22]. Exactly the context- aware computing seems to be the most important approach to realize an adaptive approach to the distribution of services and contents [32]. The Context Awareness marks the systems capable of perceiving the contexts where they and the user act, and of consequently modifying their behavior and exchanged information. These opportunities offered by the new pervasive technologies have deeply affected the world of the services providers [31].

The service used in a given context, encouraging interaction among the parties, favors the integration of resources, expertise, information and interests (before, during and after use).

The mobile devices are also enriched with multimedia contents (among which audio, video, photo files, etc.) and tools to realize these contents (among which integrated cameras, recording microphones, etc.). This allows creating and exchanging multimedia contents in real time and in any place and, therefore, realizing and sharing digital stories during the itineraries of visit. These characteristics are exploited in all the areas, included those connected to tourism and cultural heritage, in particular in the museums [13, 29]. In fact, they represent places where stories can be told and where several approaches of storytelling are tested [4, 11]: basically, the narration has evolved including digital contents to make the participation of the museums' visitors mainly interactive through a more personal and adaptive storytelling.

Finally, in order to improve the sensation of amusement for the user, it is born a storytelling where the videogames become a more effective learning tool than the traditional teaching methods, where the decisions of the player affect the global story of the play [5].

The objective of this work is to indicate an approach for the realization of a dynamic storytelling engine that can allow the dynamic supply of narrative contents, not necessarily predetermined and pertinent to the needs and the dynamic behaviors of the users.

The adaptability to the context, the social networks and the mobile world represent the pivotal points, in order to give a technological solution to the digital storytelling, which combines, indeed, the adaptive, social and mobile approaches.

The adaptive approach inserts itself in order to enhance the idea of a narrative experience able to dynamically change in relation with the elements of the profile and the context. For this purpose, it has been realized a model for the representation of all possible contexts in order to provide users of contextual services. The fundamental feature of the social component of our system is the sharing of the role of the author among many individuals, each of which, in turn, takes part into the creation of a portion or a segment of the story, but none of them can assume its paternity. Finally, we have chosen to create the system through a mobile application as it proves to be a perfect tool to have in real time all necessary information, facilitating the access to a series of information and conducting an educational role too.

In the following paragraphs, we are going to analyze the main works concerning the digital storytelling. Later, we will describe a scenery as example and introduce the general architecture of the proposed model, presenting in detail each module and phase.

The paper ends with the description of a realized application and its experimentation.

2. Related Works

In this section, we are going to show several cases of use of the Digital Storytelling, in order to make the visitors' experience more interactive and engaging, proposing personalized stories using new technologies (geolocation, mobile applications, augmented reality, etc.).

In [8] Young has proposed an interactive narrative architecture that 're-projects' the events of the story, using the planning. In the wake of the work by Young, many adaptive narrative theories have been developed, such as the use of an algorithm of user modeling to predict their actions and the trigger events, in real time; in this way, the character of the user is given by their actions with respect to the context of the story.

An example is Reading Glove [30], system of adaptive interactive narration that uses a smart 'recommender system' and a tabletop display. The reasoning engine of the system guides the users through the story, using three different modalities of recommendation: the casual ones, those based on the content of the story and the recommendations based on the user. Therefore, the recommender system operates like a kind of 'expert narrator', which follows the reader through the narration.

A further example is Framework for adaptive storytelling [10]: the story develops around the notions of interest and preferences of the user. This feedback updates the user model according to their interests through the profiling module. The story is made up of a series of correlated events and the event selection module determines what content is more suitable for the interests of the user in order to select the right event for the progression of the story.

Instead, in [2], the purpose has been to try to create an emphatic adaptive narrator. The system, under development, aims at using the emotional expressions generated by an avatar or a robot in addition to the answers of the listeners who are monitored in real time. Using several devices and sensors, the system points at codifying in runtime the listener's positive and negative emotions and on them modifying the techniques of the narrative discourse.

Then, it is possible to identify and list a very long series of 'collaborative' and 'social' storytelling that have purposes of pre-schooling, amusement, teaching and fruition of cultural heritage; among these, we would like to remember Casting [25], a software system that supports the audio-based narration of collaborative groups in the creation of non-linear stories. Casting is made up of a client and a web portal. In particular, the client allows the users to create a project team, add audio recordings, connect the audio recordings, select and publish a linear story. The client's users can recover the most up-to-date version of the story and synchronize their modifications in local. Instead, the web portal allows the users to publish the podcasts and discuss, comment, vote and reuse audio-based stories.

Moreover, the user's experience, thanks to the mobility and the territorial and thematic localization, can renew itself using the physical environment like a background and plot on which basis structuring and defining stories that wind along the urban topography [18].

The use of augmented reality is one of the services able to improve in a substantial way the interaction and the active participation of the visitor, bringing them in a more engaging way nearer to the narration of the place, superimposing the actual reality to a virtual one that stimulates the imagination and favors sense connections.

Frequency 1550 [27] is a classic example of role- play in augmented reality, presented by Waag Society and realized to allow the students to get closer to the medieval history of the city of Amsterdam. Thanks to the use of smartphone devices, the visitors move in a hybrid reality that alternates the modern and current city with the faraway one of 1550. Structured as a treasure hunt, the students move inside the city guided by the mobile device looking for the solution of the story.

Another example of cohesion of augmented reality and interactive narration with the use of mobile devices is the project AIRIS [17], developed in 2011, which puts together the concept of game with the visit of the city, calling the visitor to participate into the discovery of the place and its history through fluid and competitive modalities that amplify the curiosity and facilitate the access to knowledge heritage, connecting real space and game space. The user is guided inside a physical space that can be more or less extended: an entire city or even only one part of it, a museum, a room, a castle, etc.

In order to identify a personalized story, suitable for the needs of large masses of visitors and tourists, our work has been aimed at the definition of appropriate models and solutions of fruition that make the visit experience more appealing and immersive. In particular, there is the realization of technological solutions for the modulation of information depending on the users and their characteristics, through techniques of dynamic profiling and solutions for the generation of narrative contents coherent with the experiential mission of the visit.In fact, on the basis of the social behavior of each user, of their own interests and contextual information, it has been realized an ontological model, in order to elaborate the obtained data and propose a personalized story of the user, including a series of information about the place where they are and about the places they can visit.

3. Motivating example

In this section, we describe a typical example of use of Digital Storytelling in the field of Cultural Heritage and Tourism in order to understand to the best the main characteristics and purposes of the proposed system.

The idea is to support the tourist in order to give them, in an automatic way, a tailor-made story that contains information about the reached place and suggestions about places to visit according to their own interests, obtained through their own and their friends' likes and tags, taken from the social networks.

Therefore, the system will have to give a personalized story that includes:

- information about the visited place (main characteristics, historical news, etc.);
- specific points of interest and events for the user, filtered for category, with multimedia deepening about each point;
- contextual services for the user;
- experiences lived by other users.

For this, we want to realize a system of dynamic and contextual digital storytelling that is able to collect and elaborate information and social contents about the users giving them a personalized story about the place that they are visiting.

Let us consider now a practical example. Frank is a tourist who arrives for the first time at the harbor of Salerno, in Campania (Italy), to spend some days on holiday. Through the realized application, he will be able to obtain immediately a story about the town he wants to visit, enriched by the tourist attractions of that place. Gathering his geographical position, the system will get from the Web all necessary information and will show them to the user who will not have to worry about finding them manually. After this first phase, Frank will be able to decide to make the login using his own Facebook credentials. The system initially will calculate, according to an analysis of the likes of the same user and of his friends, all the typologies of interest and so will be able to further enrich the story through those resources whose categories correspond to the calculated user's profile. Subsequently, on the basis of an analysis of the tagged places, the system will suggest mainly those places previously visited by his friends, including in the narration their stops too. Finally, on the basis on the user's context, according to his/her current position and global profile, the system will also suggest contextual services.

At this point, Frank, through the ad hoc built story, will be able to know the history of Salerno, the points of the town he is mainly interested in, the surrounding places visited by his friends and the experiences lived by the other users, through a latest analysis that the system will make about the comments that the latter have left on TripAdvisor after their visits. Finally, when Frank, from inside the story, will select a link associated to a point of interest, he will have available multimedia contents and useful tools such as, for example, Google Maps, which will calculate on the map the itinerary to reach it, also on the basis of the means available at that moment for the user.

4. System architecture

In this section, we describe the architecture of the proposed system, which has the purpose to give the user a personalized tailor-made story, related to the place where they are.

The system architecture, shown in figure 1, is composed by three main blocks: an adaptive-social-mobile and context-aware digital storytelling module, a data management module and a knowledge base. It is made up of the adaptive-social-mobile and context- aware digital storytelling module, divided into the search, elaboration and presentation submodules, which correspond respectively to three separate phases, which are fetch, decode and execute. The search submodule has the task of collecting the data online in order to give the elaboration submodule information about the user's current position (for example, data from Wikipedia and comments of the user on TripAdvisor), the resources, services and events that are near them (e.g., a restaurant, a cinema) obtained from social networks and our knowledge base and the Facebook profile of the user and their friends. The elaboration submodule manages all received information and gives the presentation submodule the story about the town related to the user's current position, a recommended list of contextual resources, services and events and places tagged by the user and their friends on Facebook.

Finally, the presentation module has the task of proposing, as a well-structured story, all these personalized, adaptive and contextual information. The obtained final result is presented to the user through a mobile application.

The Knowledge Base is a special type of database for the management of knowledge and information: it is continuously updated by the results obtained through interaction with the application and with the Data Management Module. The Knowledge Base is composed of four components: in particular, "Users", representing all users of the application, "Services", which describes all the services of every possible application context, "Resources", which forms all the points of interest and "Events", which describes all events; and finally the Management Module (MM), used both by the administrators of the app and the users themselves. This module deals with some important issues, including: POIs and events management, where the insertion can be done directly from map, manually or by search of interests; services and comments management, interacting with TripAdvisor, Facebook and Google API.

In the following paragraph, all the submodules and the related phases of the proposed module of adaptive-social-mobile and context-aware digital storytelling will be analyzed in detail.

4.1 Adaptive-social-mobile and Context-Aware Digital Storytelling Module

The adaptive-social-mobile and context-aware digital storytelling module, presented in detail in figure 2, consists of three sequential phases: *fetch*, where information of interest are collected; *decode*, which allows their interpretation and elaboration; *execute*, thanks to which the elaborated and personalized information are proposed.

4.1.1. Fetch phase. The fetch phase is made up of three main operations: the acquisition of news about the current position, the research of the resources, services and events and the acquisition of the profile and of the user's interests.

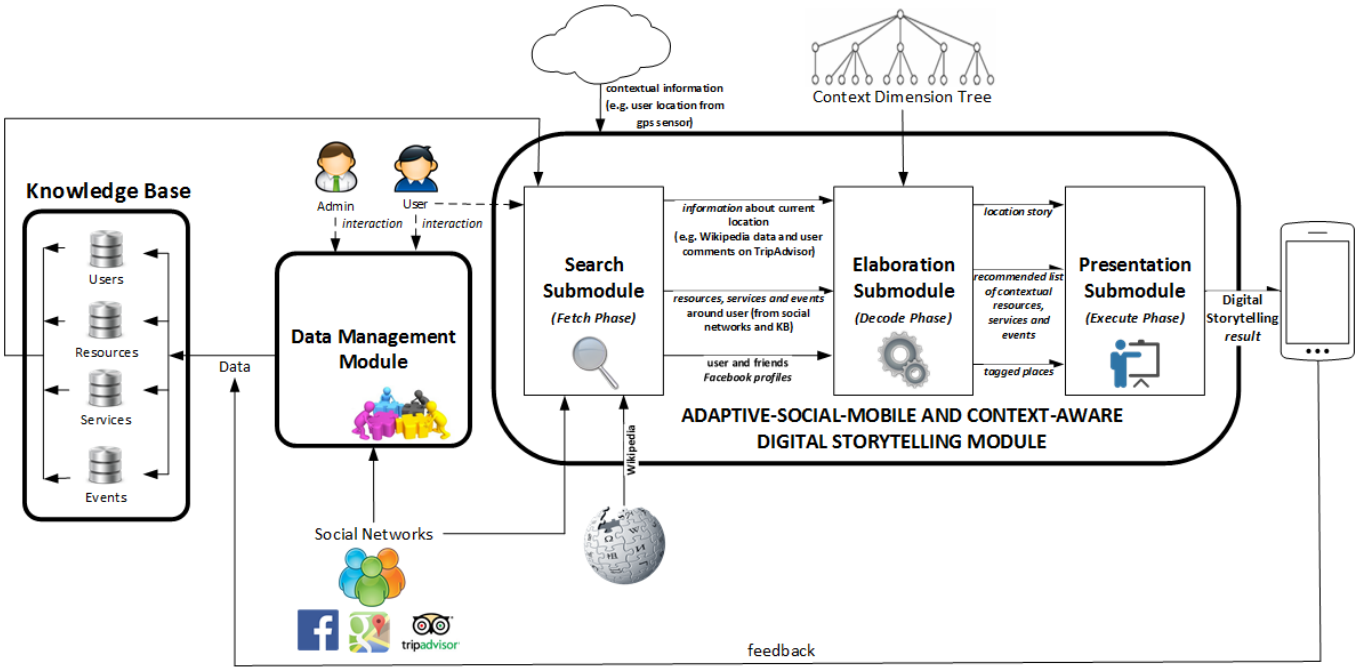


Figure 1. System Architecture

As regards the first operation, initially it is gathered the position of the user through the GPS sensor of their mobile device: it is necessary to use a service of Reverse Geocoding, that is the process of translation of the geographical coordinates in a 'human-readable' address, made available by the Google Maps API.

Then, to have the description of the locality on the basis of the gathered position, information from Wikipedia have been chosen, in particular using the MediaWiki's RESTful web service API: an easy access to services, data and metadata through HTTP requests.

Often, some coordinates of reference are associated to the pages of Wikipedia and our approach exploits this property: when you make a research of a page related to a village or a city, you obtain the nearest to the user's coordinates.

The story to present is then enriched with the experiences lived by other users, using their comments on TripAdvisor.

Then, it is made the research of the resources from social networks and from our knowledge base. In first case, model uses the map method of the TripAdvisor API, which allows looking for attractions (museums, sports structures, shops, …), hotels and restaurants through, respectively, the map attractions, map hotels and map restaurant methods. In case the user does not have a social account, they can get the main points of interest of the places where they are using the location-mapper method of the TripAdvisor API. In second case, research is done from our knowledge base, continuously fed by users and results previously obtained from the web.

Finally, for the personalization of information, it is exploited the third operation of the fetch phase of our module, whose objective is to access the social data of the user to get, for example, their profile, their likes, their events and preferred places, their friends: this happens by querying the knowledge base and through the use of the Facebook API and the analysis of some principal methods, such as user likes, user friends, user events and tagged places.

4.1.2. Decode phase. The decode phase is formed by data elaboration and Context Dimension Tree elaboration. The first operation consists in the elaboration of data, in particular in the management of information get from Facebook, in the following correspondence with those of TripAdvisor and in the calculation for the proposition to the user of the resources, that is, of the geographical points of interest according to their own preferences.

In particular, to determine the user's interests, as well as the categories to recommend, we consider their most recent likes on the pages of the Social Network: to any like it corresponds a specific subcategory of Facebook that our system associates to a category of TripAdvisor; for example, to a like on a page related to an art gallery or a painter it will correspond an occurrence of the museum category.

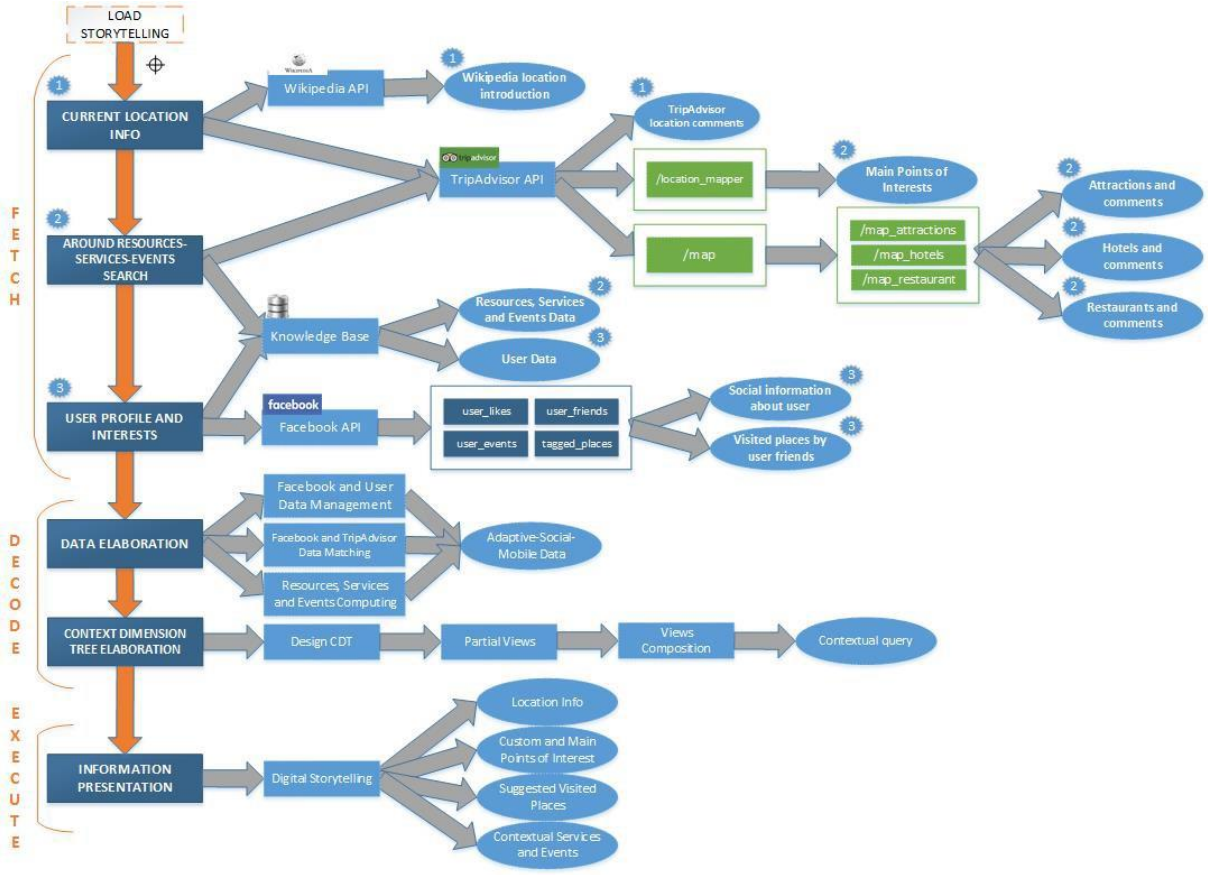


Figure 2. Adaptive-social-mobile and Context-Aware Digital Storytelling

Through our system, it is possible to calculate the occurrences and get the preferred categories of interest; then, on the basis of the contextual approach, resources, services and events will be proposed to the user, within the personalized story.

In fact, the resources are properly filtered and contextualized based on the second operation of the decode phase: Context Dimension Tree elaboration.

A key element in the design of a contextual application and a Context-Aware System is the representation and management of the context itself: this operation is made through the Context Dimension Tree [16].

CDT is a tree composed of a triad <r; N; A> where r indicates its root, N is the set of nodes of which it is made of and A is the set of arcs joining these nodes.

CDT is used to be able to represent, in a graphic form, all possible contexts that you may have within an application.

Nodes present within CDT are divided into two categories, namely dimension nodes and concept nodes. A dimension node, which is graphically represented by the color black, is a node that describes a possible dimension of the application domain; a concept node, on the other hand, is depicted by the color white and represents one of the possible values that a dimension may assume. Each node is identified through its type and a label.

The children of the root node r are all dimension nodes, they are called top dimension and for each of them there may be a sub-tree. Leaf nodes, instead, must be concept nodes. A dimension node can have, as children, only concept nodes and, similarly, a concept node can have, as children, only dimension nodes.

In addition to nodes, you can use other elements: the parameters, which may arise both from a dimension node (graphically represented by a white square) and from a concept node (white triangle), submitting them to particular constraints. In fact, a concept node can have more than one parameter, while a dimension node can have only a parameter and only in case it has not already children nodes. The introduction of parameters is due to their usefulness in shaping the characteristics that can have an infinite or very high number of attributes. For example, a node representing Cost dimension risks having a high number of values that should be specified by as many concept children nodes. In a similar case, it is therefore preferred to use only one parameter, whose value will be specified in each case. Leaf nodes, in addition to concept nodes, can also be parameters.

In general, each node has a parameter corresponding to a domain, dom(nP). For parameter nodes connected to concept nodes, the domain can be a set of key values from a relational database, while in case of parameter nodes connected to dimension nodes, the domain is a set of possible concept nodes of dimension.

In figure 3, it is shown a general designed CDT, called Meta CDT, which is the starting point for the design of a specific CDT that can be exploited in contextual applications.

You may note six top dimensions, which correspond to the questions of the 5W1H method: Location (WHERE), Role (WHO), Time (WHEN), Situation (HOW), Interests (WHAT) and Utilization (WHY).

In particular, there are two types of users and eleven categories of interests.

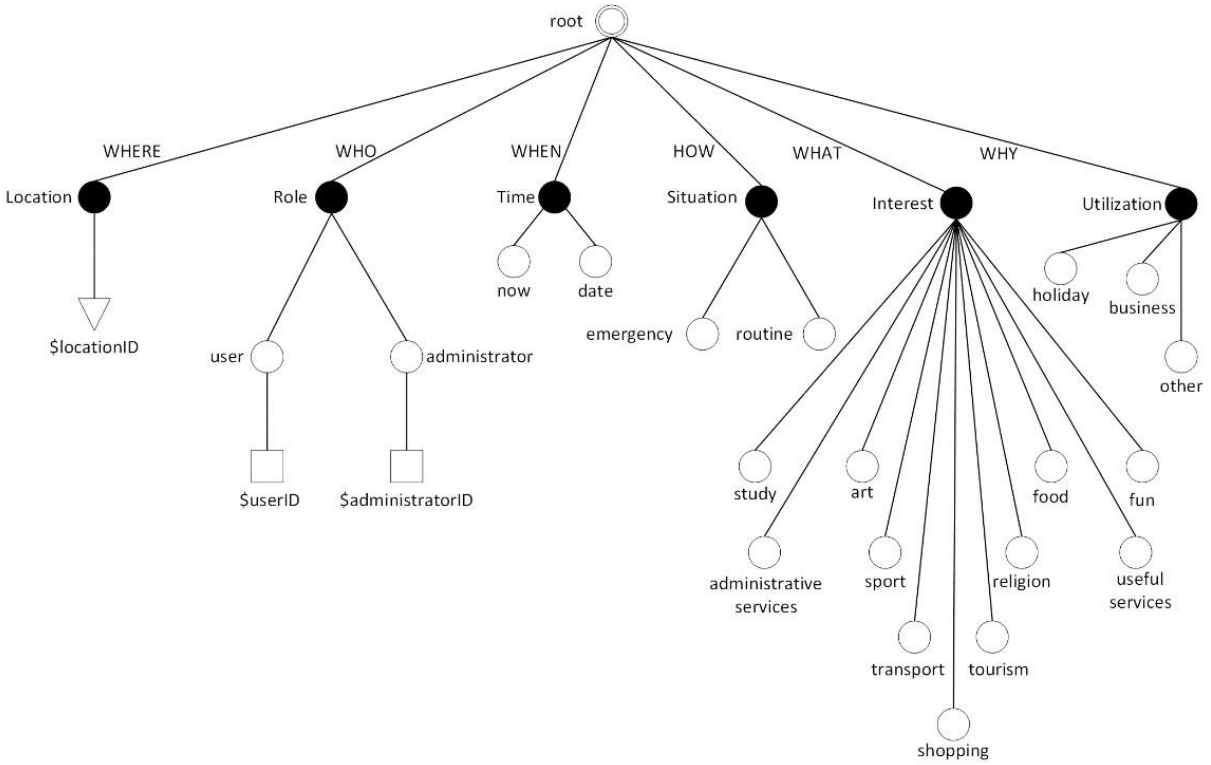


Figure 3. Meta CDT for contextual applications

A context element is defined as an assignment d_namei = value, where d_namei indicates a possible size or undersize of CDT (it is the label of a dimension node), while value may represent the label of one of the concept nodes that are children of the considered dimension node or the value of a parameter referring to one of these concept nodes or the value of a parameter referring to the considered dimension node. For example, these assignments are possible context elements: Interest = tourism, Location = LocationID (ID = 3), Role = user, Utilization = holiday.

A context is specified as: $\wedge$ *(d_namei = value).*

It is defined as an "and" among different context elements.

Several context elements, combined with each other by means of an "and", damage, therefore, the origin of a context.

For example, a possible framework that can be obtained from the previously seen CDT, through the context element that we have listed, is:

C = (Location = locationID (ID=3)) $\wedge$ *(Role= user (ID=15))* $\wedge$ *(Time = now)* $\wedge$ *(Situation = routine)* $\wedge$ *(Interest = tourism)* $\wedge$ *(Utilization = holiday)*

The context is defined as a user, interested in tourism, who uses the contextual app on vacation, in a called place.

Therefore, through the Context Dimension Tree, it is possible, after analyzing the domain of application, to express the size characteristics and values they can take in a graphical way by, respectively, dimension nodes and concept nodes or parameters.

The assignment to a dimension of one of its possible values is a context element. The context element can be considered the main feature of the application, by which a context can be decomposed. The moment you make the formulation of the context, you must specify all the context elements that are part of it and that enable its creation. Any context is expressible by an "and" combination of the context elements to which they are peculiar.

By definition, you can begin to understand how you will create views based on data relating to each context; in fact, they will be built starting from the portions of the database and then from the partial views, associated to the context element that takes part into context information.

The Context Dimension Tree elaboration is composed of methodologies and phases to obtain contextual service.

The methodology, shown in figure 4, has been realized in order to manage the database and to carry out reductions of their content based on the context.

The purpose is to help the designer in the definition of all contexts relevant to the considered application and, later, in the association to each context of the portion of the database containing the relevant data about the context. The methodology consists of three main phases, which we will see in detail later: design phase of the Context Dimension Tree (CDT), definition phase of partial views and composition phase of global views.

1. Design phase of the Context Tree: in this phase, the Context Dimension Tree is designed to identify significant context elements for the considered application. In fact, it focuses on the definition of contexts and on the elements that compose them. These contexts must be identified and shaped, indicating particular elements that characterize each of them. As it has been said, it is available a special tool called Context Dimension Tree (CDT) to make context design. Three CDT were made for specific environments in order to represent and manage a multitude of different contexts and in order to identify, represent, preserve and make available cultural points for each type of user.

2. Definition phase of partial views: after the definition of all the contexts and their context elements, in this step a different portion of the database is associated to each context element, containing the relevant data for it.

In practice, the goal is to find the appropriate value for a given dimension, in order to obtain, by means of the values of all the dimensions, a valid query and specific to the context in which the user is located.

A partial view could be related to dimension "Role": once logged in, the application is able to recognize the user and to know more precisely whether he/she is, for example in tourist areas, a resident or a tourist. Thus, the value "tourist" of dimension "Role" is a partial view for the current context: using this knowledge, you can exclude certain services, not suitable or useful to the tourist role.

3. Composition phase of global views: this is the phase where you have the automatic generation of views associated with each context, which is made starting from partial views associated with context elements. After the creation of the global views of the contexts, the answers to questions that will be asked to the system will be developed from these views and, in particular, from the view associated with the context in which you are located when the query is performed.

In particular, once defined the values for each dimension, you can use all the information obtained in order to identify the right context and offer data and services customized for the user.

It is assumed the example of a tourist who is walking near a beach who gets initially a notification of his/her proximity. Later, he/she needs to deepen such notification. Therefore, it will propose him/her services that they might be interested in, such as the site of the nearest beach, where he/she can get the price list.

4.1.3. Execute phase. The third phase of the process, execute, is based on the presentation of contextualized and personalized information: therefore, the result shows itself like a well-organized story that presents a general introduction about the place reached by the user, enriched with the experiences shared by other users, a list of the main suggested attractions and information about the near places visited by the friends.

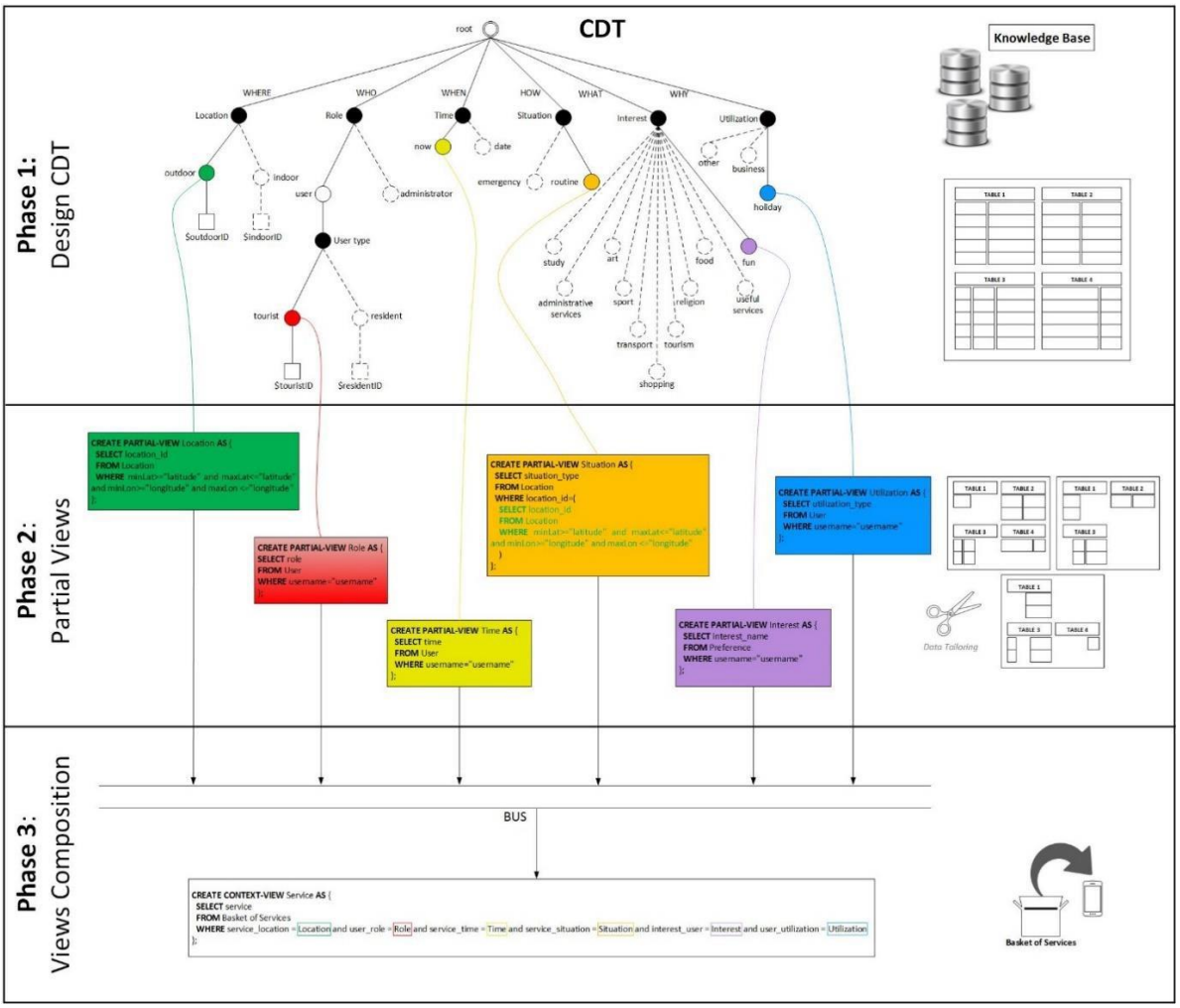


Figure 4. General System Workflow

Now, consider an example about a user called Matthew who puts some likes on social pages associated to typologies of interests: for example, school, university, library, pizza and Italian food. Each typology of interest represents a category: culture for school, university and library, restaurant for pizza and Italian food.

On the basis on this information and the data obtained on the Knowledge Base and on the web, our system determines the user's context, identifying the most appropriate resources, services and events, on the basis primarily of the profile and the user's location.

5. An adaptive Digital Storytelling App

In this section, we are going to present the planned and implemented app according to what previously described. The system is made available to the user through a hybrid mobile application, which allows an high level of personalization proposing a digital story that includes a series of tailor-made information, among which the suggestion of places of interest according to their current position and their own social profile.

When the user starts the app, if they do not have a Facebook account or have not logged in yet, they receive some information about the place where they are and the main points of interest that they can see on the map and reach through a navigator on the device. Once effectuated the login, the system proposes to the user a personalized story, which includes, this time, in addition to information about the place, also points of interest tailor-made suggested to the user, depending on their interests get form Facebook, and suggestions about near places that they have visited. For each of these elements it is possible to visualize the rating, the address, the distance from the current position, the preview of the users' last comments and the related web page of TripAdvisor. Furthermore, user has available contextualized services and events. Finally, if some of the user's Facebook friends have been near the current place, the system communicates when and what they have visited.

The screenshots of the application are presented in figure 5, the platform of Data Management Module is shown in figure 6.

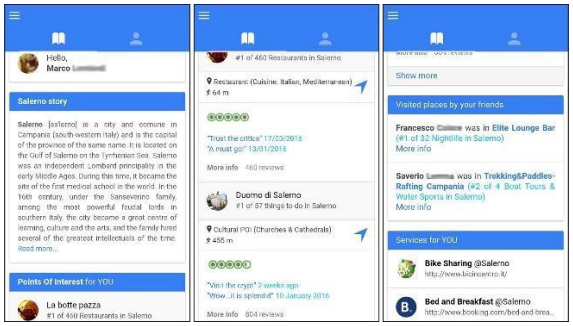


Figure 5. Screenshots of the application

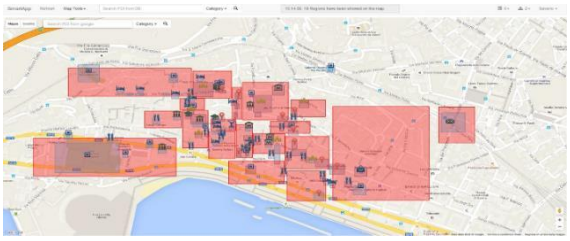


Figure 6. Platform of Data Management Module

5.1. Experimental Results

Now, we are going to present the conducted analysis, using the realized App, to understand the influence on the user experience of the proposed digital storytelling model.

The 40 participants, recruited by invitation, are above all students or university professors and do not know the subject of the study. All live in different towns, in Italy or abroad, are between 18 and 57 years old, are registered to Facebook and have an Android or Apple mobile device. The experimental phase is based on the storytelling model evaluation. Initially, the App has been distributed and installed by all the participants, to whom it has been requested to log in with Facebook credentials.

After having interacted for some days with the application, the participants have then answered on the basis of the Likert scale to seven statements: (1) the system gives right information about the narration of the place; (2) the system has included in my profile the categories I am mainly interested in; (3) it has proposed me points of interest and events I am mainly interested in; (4) the experience of friends and other users has been useful for me; (5) the system has suggested me tailored services; (6) the system is immediate and user-friendly; (7) the narration has been exhaustive through a personalized guide.

Table 1 presents a synthesis of the answers of the participants to each declaration.

Table 1. Experimental results

Answer	1	2	3	4	5	6	7
Strongly disagree	0	1	2	0	1	1	2
Disagree	2	3	5	2	3	2	2
Neither agree nor disagree	5	4	5	3	9	6	4
Agree	19	16	16	14	13	12	18
Strongly agree	14	16	12	21	14	19	14

As shown in this table, of the 40 participants who have interacted with the application, many agree and/or strongly agree that the system gives appropriate information about the place, supplies the resources of interest depending on their preferences and the final narration results complete and pertinent. Instead, only in few cases, the participants do not are particularly satisfied.

6. Conclusions

In this paper, it has been introduced an original approach to digital storytelling. In particular, the proposed system is based on the adaptive, social and mobile approaches, using a contextual model with the purpose to create a dynamic digital storytelling system able to collect and elaborate information and social contents about the user giving them a personalized story concerning the place they are visiting. It has been studied a real case developing a mobile application. The obtained results of the experimental campaign are satisfying and show the good potentialities of this kind of approach. The future developers can predict a greater integration of the system with new heterogeneous information sources, in order to obtain a more and more detailed and personalized final story.

7. References

[1] Alexander, B. and Levine, A., 2008. Web 2.0 Storytelling: Emergence of a New Genre. *Educause Review*, November-December 2008, 40-56, Boulder.

[2] Bae B. C., Brunete A., Malik U., Dimara E., Jermsurawong J. and Mavridis N, 2012, July. Towards an Empathizing and Adaptive Storyteller System. *In Eighth Artificial Intelligence and Interactive Digital Entertainment Conference*.

[3] Barthes, R. Introduction to the Structural Analysis of the Narrative. Occasional Paper, Centre for Contemporary Cultural Studies, University of Birmingham, 1966.

[4] Bedford, L., 2001. Storytelling: the real work of museums. *Curator: the museum journal* 44.1 (2001): 27-34.

[5] Champagnat R., Delmas G., Augeraud, M., 2012. A Storytelling Model for Educational Games. *Proceedings of the 6th European Conference on Games Based Learning*.

[6] Colace, F., De Santo, M., Greco, L., Lemma, S., Lombardi, M., Moscato, V. and Picariello, A., 2014. Context-Aware Framework for Cultural Heritage Applications. *10th Signal-Image Technology and Internet-Based Systems (SITIS).*

[7] Chang, S.K., Yung, D., Colace, F., Greco, L., Lemma, S. and Lombardi, M., 2015. An Adaptive Contextual Recommender System: a Slow Intelligence Perspective. *27th Software Engineering and Knowledge Engineering (SEKE)*.

[8] El-Nasr, M.S., 2007. Interaction, narrative, and drama: Creating an adaptive interactive narrative using performance arts theories. *Interaction Studies*, 8(2), 209-240.

[9] Gabriel, Y. and Connell, N.C., 2010. Co-creating stories: Collaborative experiments in storytelling. *Management Learning*, 41(5), 507-523.

[10] Garber-Barron, M. and Si, M., 2012, July. Towards Interest and Engagement: A Framework For Adaptive Storytelling. *In Eighth Artificial Intelligence and Interactive*

Digital Entertainment Conference.

[11] Johnsson, E., 2006. Telling tales. A guide to developing effective storytelling programmes for museums. *Renaissance London. C. Adler. London, London Museums Hub*.

[12] Kickmeier-Rust, M.D., Göbel, S., and Albert, D., 2008. 80Days: Melding adaptive educational technology and adaptive and interactive storytelling in digital educational games. *In Proceedings of the First International Workshop on Story-Telling and Educational Games (STEG'08)*.

[13] Ioannidis, Y., El Raheb, K., Toli, E., Katifori, A., Boile, M. and Mazura, M., 2013. One Object many stories: Introducing ICT in museums and collections through digital storytelling. *Digital Heritage International Congress*.

[14] Lombardo, V., and Damiano, R., 2012. Storytelling on mobile devices for cultural heritage. *New Review of Hypermedia and Multimedia*, 18(1-2), 11-35.

[15] Malkawi, R., 2012. Integration of multi-media technologies to facilitate reflection and learning, particularly in the area of digital storytelling. *International Conference on Education and e-Learning Innovations (ICEELI)*.

[16] Tanca, L., Bolchini, C., Curino, C., Schreiber, F.A., 2006. Context integration for mobile data tailoring. Italian Symposium on Database Systems (SEBD), pp. 48-55.

[17] Nisi, V., Oakley, I. and Haahr, M., 2008. Location- aware multimedia stories: turning spaces into places. *Universidade Catolica Portuguesa*, 72-93.

[18] Paay, J., Kjeldskov, J., Christensen, A., Ibsen, A., Jensen, D., Nielsen, G. and Vutborg, R., 2008. Location-based storytelling in the urban environment. *In Proceedings of the 20th Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat* (OZCHI '08). ACM, New York, NY, USA, 122-129. DOI=http://dx.doi.org/10.1145/1517744.1517786

[19] Pearce, C., 1994. The ins & outs of non-linear storytelling. *SIGGRAPH Comput. Graph.* 28, 2 (May 1994), 100-101. DOI=10.1145/178951.178956 http://doi.acm.org/10.1145/178951.178956

[20] Pittarello, F., 2006. A Simple Story: Using an Agents' Based Context-Aware Architecture for Storytelling. *Technologies for interactive Digital Storytelling and Entertainment* (pp 25-36).

[21] Colace, F., De Santo, M., Moscato, V., Picariello, A., Schreiber, F.A., Tanca, L., 2015e. *Pervasive Systems Architecture and the Main Related Technologies*. Data Management in Pervasive Systems, pp. 19-42.

[22] Colace, F., De Santo, M., Moscato, V., Picariello, A., Schreiber, F.A., Tanca, L., 2015d. *PATCH: A Portable Context-Aware ATlas for Browsing Cultural Heritage*. Data Management in Pervasive Systems, pp. 345-361.

[23] Robin, B., 2006. The educational uses of digital storytelling. In C. Crawford et al. (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference*, (pp. 709-716). Chesapeake, VA: AACE.

[24] Robin, B., 2008. Digital Storytelling: A Powerful Technology Tool for the 21st Century Classroom. *Theory Into Practice*, v47 n3, pp 220-228.

[25] Schumann J., Buttler T. and Lukosch S., 2010. Supporting asynchronous workspace awareness by visualizing the story evolution in collaborative storytelling. *In Collaboration and Technology* (pp. 218-232). Springer Berlin Heidelberg. A Collaborative technology to enhance field trips.

[26] Scolari, C.A., 2009. Transmedia storytelling: Implicit consumers, narrative worlds, and branding in contemporary media production. *International Journal of Communication*.

[27] Spallazzo D., 2013. *Davide Spallazzo - Musei, fruizione culturale e tecnologie*. Storicamente, Vol. 9, ISSN:1825-411X.

[28] Spaniol, M., Klamma, R., Sharda, N. and Jarke M., 2006. Web-based learning with non-linear multimedia stories. *In Advances in Web Based Learning–ICWL*, (pp. 249-263). Springer Berlin Heidelberg.

[29] Springer, J., Borst Brazas, J. and Kajder, S., 2004. Digital storytelling at the National Gallery of Art. *Museums & the Web*. D. Bearman and J. Trant. Arlington, VA, Archives & Museums Informatics.

[30] Tanenbaum, K., Hatala, M. and Tanenbaum, J., 2011. User perceptions of adaptivity in an interactive narrative. In *Proceedings of the 19th international conference on User modeling, adaption, and personalization* (UMAP'11), Joseph A. Konstan, Ricardo Conejo, José L. Marzo, and Nuria Oliver (Eds.). Springer-Verlag, Berlin, Heidelberg, 389-400.

[31] Colace, F., De Santo, M., Greco, L., 2014. *An adaptive product configurator based on slow intelligence approach*. International Journal of Metadata, Semantic and Ontologies (IJMSO), Vol.9, No.2, pp. 128-137.

[32] Colace, F., De Santo, M., Greco, L., Moscato, V., Picariello, A., 2015b. *A collaborative user-centered framework for recommending items in Online Social Networks*. Computers in Human Behavior 51: pp. 694- 704.

[33] Colace, F., Moscato, V., Quintarelli, E., Rabosio, E., Tanca, L., 2015c. *Context awareness in pervasive information management*. Data Management in Pervasive Systems, pp. 235- 256.

Scaffolding Version Control into the Computer Science Curriculum

Denise M. Case, Nathan W. Eloe
School of Computer Science and
Information Systems
Northwest Missouri State University
Maryville, MO 64468, USA
{dcase,nathane}@nwmissouri.edu

Jennifer L. Leopold
Department of Computer Science
Missouri University of Science and Technology
Rolla, MO 64468, USA
leopoldj@mst.edu

Abstract

Version control systems (VCS) are widely-used in the software industry. They provide a powerful, collaborative framework that allows software engineers to work together effectively. VCS allow users to track changes and merge ongoing work into concurrently evolving software projects. Distributed VCS such as Git, allow a great degree of flexibility, and provide powerful options for managing personal code and evolving collaborative content. Power incurs responsibility, and introducing collaborative coding and version control tools to new developers can create many challenges. Yet these tools, once mastered, are crucial skills for professional developers. In this paper, the authors introduce VCS to computer science students both in a custom environment specifically designed to support new developers and in a commercially-available native environment suitable for more experienced students. Results show that proper introduction of these powerful tools can make early exposure a positive and valued experience.

Keywords. *collaborative systems, computer science education, education technology, pedagogy, profession-based learning, scaffolding, version control, visual language*

1 Introduction

Computer science courses aim to prepare students to think computationally and *apply* their thinking in a practical way. Many powerful tools and techniques are being developed in both academia and industry to assist with the practical application of computer science and software engineering. Incorporating a variety of these useful elements into a crowded curriculum can be challenging, but also increasingly valuable, both for the benefits it offers graduates in early employment, as well as in the way these tools and techniques can be used to enhance the immediate educational experience. One tool that has developed to be nearly ubiquitous in industry is the application of advanced *version control systems* (VCS). In industry and in its prequel, *project-based learning*, development projects require collaboration between teams of developers to effectively and efficiently create software that addresses real-world needs. Application of VCS and the associated mastery of these systems facilitates teamwork, collaboration, and effective and enjoyable creation of significant and useful software.

VCS and other professional tools are often introduced later in the educational curriculum, possibly in a software engineering course, generally a second-year or later course, and may not be reinforced through repetition and application in later classes. This is unfortunate, because powerful tools such as VCS can also be difficult for new users and especially those new to programming and software in general. The power and flexibility of new distributed VCS can make them especially challenging, unfriendly, and distracting for students already fully engaged in developing core competencies required in the computer science curriculum.

This paper, in conjunction with concurrent research [1], is an investigation into better approaches to introduce version control into pedagogy. For novice, less experienced students, we propose introducing these powerful version control tools by providing facilitative scaffolding in the form of a simple, consistent, visual language, that supports students during their introduction and initial exposure to VCS. For more advanced students, who may still be entirely or mostly unfamiliar with VCS, we propose a series of short, simple tasks to progressively integrate version control while working in a native environment.

2 Background and Related Work

In this paper, we build on research in instructional technology and pedagogy, in visual languages, in collaborative frameworks, systems, and learning, and on prior work on

DOI reference number: 10.18293/DMS2016-036

the introduction of version control systems in the computer science curriculum.

2.1 Instructional Technology and Pedagogy

Software engineering and development can be challenging; a variety of pedagogical approaches have been developed to assist with the learning process [2]. One popular technique is called *scaffolding*, designed specifically to support novices as they increasingly master challenging tasks [3]. As the novice gains proficiency, scaffolding supports are progressively removed through a process called *fading*. This continues until the supports are no longer needed. In this paper, we propose approaches for assisting students as they gain exposure and mastery with VCS and propose two approaches that allow VCS to be efficiently and effectively integrated into the learning process.

2.2 Visual Languages

Visual languages continue to grow in importance; the success of visual languages is evident in the proliferation of iconic-based languages on smart phones [4]. Such systems need to be able to reach non-expert users as intuitively as possible, and as multimedia applications continue to develop and expand, we expect visual language systems, both special- and general-purpose, to continue to grow in importance, both as the subjects of theoretical research and in actual practice [5].

2.3 Collaborative Frameworks, Systems, and Learning

Current trends in education include the integration of online, hybrid, and collaborative learning environments [6, 7]. Modern collaborative environments in higher education are responding to many of the same opportunities as industry, including the increasing application of smart technologies, such as cloud computing for software distribution and usage and the adoption of collaborative paradigms and environments that leverage the work of student teams just as they do software development teams in industry [6]. In profession-based learning and outcomes-based education, collaboration is both a teaching strategy and a learning outcome; in both public institutions and large organizations, the ability to work together is critical to successful software engineering [8, 9]. The application of Computer Supported Collaborative Learning (CSCL) to software engineering has been the subject of recent research, including versioning systems with positive results [10].

2.4 Version Control in Education

For over ten years, educators have employed a variety of VCS to promote collaboration and support student engagement in the classroom [11] and Git, specifically, has been the focus of current research and used to help students work collaboratively and receive instructor feedback [12, 13].

Widely-employed in professional software development, these powerful systems offer considerable benefits to educators, but adoption continues to suffer from the perception that VCS are difficult, time-consuming, or could distract from the core content of the course [14, 15]. Recent research recommends VCS integration *throughout the curriculum*, in a *continuous and gradual manner* [15].

Hosting solutions such as Bitbucket and GitHub offer special packages for students that include unlimited free private repositories and other benefits. To support educational users, GitHub created GitHub Classroom, a set of utilities on top of its hosting service that automates the creation of repositories and associated permissions. GitHub Classroom supports teachers by allowing projects to be assigned in various ways with various degrees of visibility to other students in the class. It integrates with Continuous Integration tools, allowing unit tests to be run on assignment submission and enables *auto-grading* in a consistent manner. Use of Bitbucket, GitHub, and GitHub Classroom requires students to create and use their own student accounts.

3 GitSubmit: A Custom Environment for Introducing VCS to Novice Developers

GitHub Education and GitHub Classroom are powerful, well-designed tools that enable the instructor to automate the creation of assignment repositories and provide feedback to students. However, these tools do not simplify students' interaction with the submission system; students must use their tool of choice to interact with the system. This can be overwhelming to students just learning to program and develop software. Developing software to solve problems requires a form of thinking foreign to many students of programming, and they are constantly being introduced to new tools such as IDEs and compilers. Submission systems designed for other kinds of assignments (such as essays or worksheets) are not ideal, but are used because they are in place and familiar to the students.

GitSubmit is an attempt to provide an interface to students that is simple, requires little to no interaction from an "administrative" standpoint, and is secure (and preserves academic integrity). Unless a student wants to further interact with GitLab, the only required interaction is changing their account password from the default. GitSubmit is built around a self-hosted GitLab [16], an open source Git

hosting solution that allows accounts to be created with the students' campus usernames during course creation.

3.1 Student Interaction with GitSubmit

GitSubmit does not require the student to install any software (including Git); it is a packaged Python executable that contains all required functionality, including ssh, authentication, and Git commands. After being greeted on first load by a simple configuration window (Figure 1) where they specify their username, password, and a place to put all of the assignments it checks out, the student sees the assignment window as shown in Figure 2. The students can select an assignment to work on by selecting the semester and class from the tree-based menu on the left hand side of the application. They begin a project by selecting the assignment, then clicking the button circled in Figure 2 (which maps to a combination of `git pull` and `git clone`, depending on whether the assignment has already been started). Upon doing so, they see the project has been downloaded as in Figure 3.

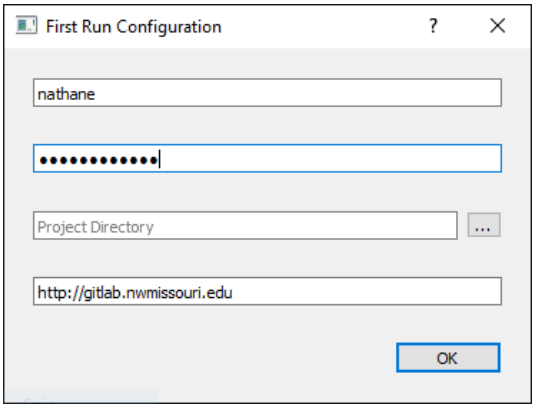


Figure 1: Configuring Gitsubmit

A student working on a project writes their code in the project directory, selects modified and new files in the "Unstaged Changes" window, then clicks the arrow between the Unstaged and Staged Changes windows (see Figure 4). This button maps to `git add`. After all required changes are added to the Staged Changes pane, the student writes a brief commit message in the box, then selects the right-most arrow (`git commit`) to commit their code. At this point, the student sees something similar to Figure 5; all that remains is to push the upload button (circled in Figure 5), which will perform a `git push` operation, and sync the code to the central repository. The green checkmark icon in the commit window will appear next to the top-most commit, signifying that it is the version of the code that will be accessed by the grader (or whoever does the next `pull`, such as a partner).

To improve security and streamline the submission process, the student's password is not saved; it is used only once when the program is first run. The password is used to obtain a token that allows the interface to authenticate to the GitLab API as the user without a password. Additionally, when a student logs in for a first time from a computer, a new SSH key is generated and used for all Git operations, specifically `push`, `pull`, and `clone`. In this way, no passwords are saved, and if the student loses their computer, access can be revoked by removing the SSH key from their account.

3.2 Instructor/TA Interaction with GitSubmit

Just as the students have a simplified interface for interacting with the submission system, the instructor interface is designed to automate much of the tediousness behind creating classes and assignments. The instructors, TAs, and graders are required to install Git and have it on their `PATH` so the interface can access it; this is because there are certain advanced features that are much more reliable using the standard Git client. The added setup complexity of the instructor's interface is worth it for the powerful features GitSubmit provides.

When the teacher wants to start a new class, they open the instructor interface, where they must provide information such as the semester, year, and a CSV containing a course roster. The instructor interface uses the CSV to ensure that all students have accounts on the system; if a student in the class does not exist in GitLab, the student's user is created using their email address as their username and a default password. The interface creates a GitLab group with teacher as the owner, and a specially formatted group name that will be used to populate the lists of courses and semesters for both instructors and students. Students are *not* added to this group, as it would allow them to view each other's assignments, but graders and TAs are added to the group so they have access to every students' submissions. The course roster is stored as a Git repository within the group.

To add an assignment, the teacher navigates to the "Add Assignment" tab, selects the course from the list (which is populated by using the list of groups the teacher is an owner of), provides an assignment name, and either a project description (in Markdown), or a skeleton directory (which is useful for distributing tests or boilerplate code). Optionally, they may also provide a comma-separated values (CSV) file containing groups. The interface creates a repository for each user (or group) with the name format `<lname><finitial>_<assignment_name>` (example: `eloen_lab11`) on the GitLab server, adds the student as a developer on the repository (so they do not have permission to change visibility to other students), and

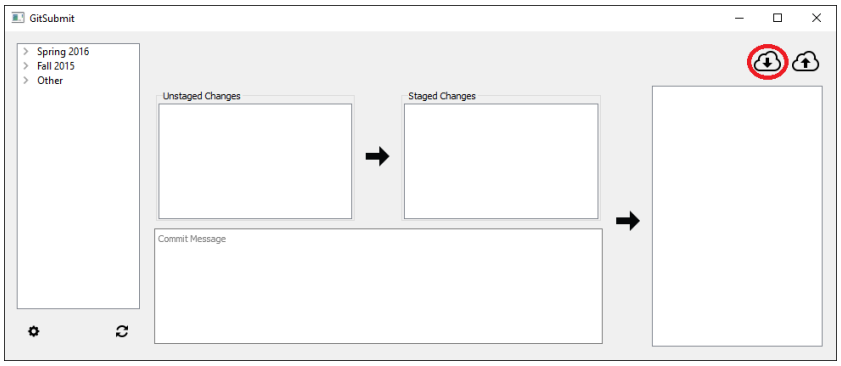


Figure 2: The Gitsubmit main window (project download button circled in red)

Figure 3: A successfully downloaded project

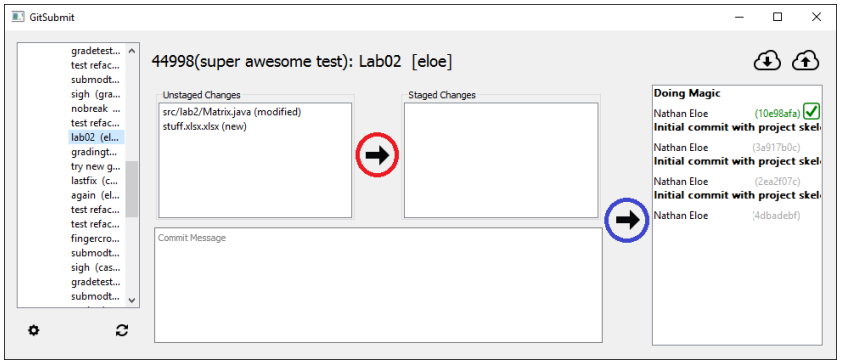


Figure 4: Committing your code.

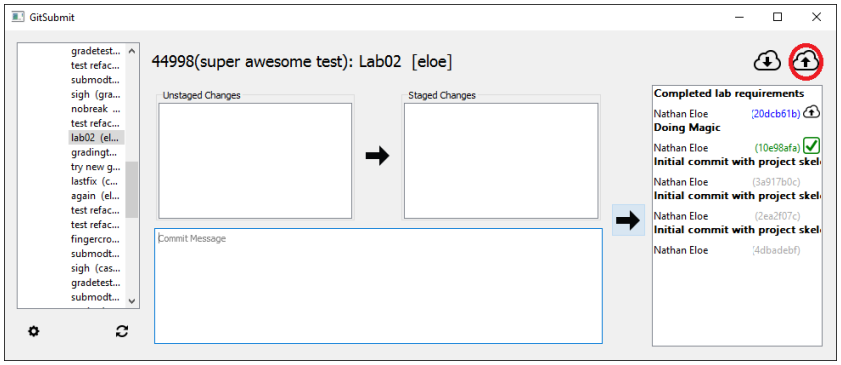


Figure 5: Finishing up.

pushes the base repository (skeleton directory or description `.md`) to GitLab. Additionally, the interface creates a single repository containing all student assignment repositories as Git submodules (for more information on Git submodules, see Chapter 7 Section 11 in [17]). This enables whoever is grading an assignment to clone a single repository, then get the newest version of all submissions by running `git submodule update --init --recursive; git submodule foreach git pull origin master`.

3.3 Submission Workflow in GitSubmit

The interactions above correspond directly to the normal academic submission process; to assign work to students, a teacher must write the assignment description, distribute it to students, and collect and grade the final submissions. Using GitSubmit, the instructor formats the assignment description (or project skeleton), and distributes it to students using the "Add Assignment" functionality of the instructor interface. Students then use the GitSubmit student interface to retrieve the assignment description and complete the assignment. A student can then submit his or her work by pushing to the Gitlab server (most submissions require either uploading to a digital dropbox or handing physical paper to the instructor). Collection of the submissions is done using the `git submodule` commands to pull the latest versions of each student's work. The individual grading the work can then grade as they normally would. Feedback beyond a final grade can be provided to students using Gitlab's comment infrastructure, which allows people to leave comments and questions associated with individual lines of code (a level of clarity one can not always achieve when using traditional learning management system (LMS) submission capabilities).

GitSubmit has the potential to vastly streamline the process of assignment submission beyond what has been implemented thus far. Because collecting submissions can be done with two simple commands, enforcing strict deadlines can be done using a task scheduler or a cron job. Additionally, instructors can take advantage of Gitlab's webhooks (and Git's pre- and post-commit hooks) to automate objective evaluation of assignments using a Continuous Integration (CI) server. Automated application of CI, static code analysis tools, and automated testing can help identify incorrectness and assess a variety of "soft" programming elements, such as adherence to coding standards. Any time savings due to such automation may allow graders to spend more time reviewing submissions at a higher level of judgement and which allows graders to spend more time evaluating student submissions on higher levels of code quality and enable teachers and teaching assistants to spend more time identifying areas for individual assistance and working with struggling students.

4 Native Git: A Gentle Introduction for Upper Level Courses

Higher-level students may be several years past their bachelor's degrees with several years of work experience. However, it is not unusual for these students to primarily be working as independent developers and have little or no experience with version control systems.

The goal for these more independent and experienced students is to introduce them to version control in a positive way and encourage them to incorporate it into their personal workflow. If we are successful, the hope is that the students will independently choose native solutions for subsequent projects and be better prepared for industry opportunities that involve coding on a professional team.

To this end, Git was introduced over a series of assignments by integrating version control into a series of assignments in an upper-level course focusing on ASP.NET Core. Bitbucket was selected as the native hosting platform because its student program at the time was the only one to offer free, unlimited private repositories for students and faculty, a feature the authors have greatly appreciated and

employed.

4.1 Getting Started

The initial assignment was conducted in a shared discussion forum, *Get Started with Bitbucket*.

1. Go to the Bitbucket Cloud Documentation Home [19].
2. Under *Get started*, click *Create an account* and follow the steps to create a personal account with your university email.
3. Under *Get started*, click *Set up version control* and follow the steps to *Set up Git* and *Install Git for Windows* (use the installation defaults). In Step 4, configure your user.name and in Step 5, configure your email.
4. You may also do Step 2 if you like; it is helpful, but not required.
5. When successful, reply with a subject of *Team x - Your Name - Success* and how long the assignment took.
6. Include a screenshot of a git bash command window running on your laptop.

The otherwise unsupported assignment was consistently completed independently and in under 10 minutes.

4.2 Personal Repositories

To introduce students to personal software repositories, students created a repository and pushed their final assignment submission code to their new repository.

1. Complete the assigned project.
2. Create a new private repository with the same name.
3. Commit and push your code up to your private repository.

Again, the additional time required was very little—and aside from simple directions, little lecture or other classroom time was required. At this time, only Git was required (most students used Git for Windows). TortoiseGit was recommended, but students were encouraged to explore (and report on) their experiences with other tools. Students reported positive experiences with SourceTree, Git Bash (command line), and the integrated Git in Microsoft's Visual Studio.

Git was incorporated into several other personal assignments and the students were assigned a discussion forum to comment on and share their experiences. Students shared resources and worked together to learn, teach, educate, and persuade others to explore various approaches.

4.3 Full Collaboration

Version control is useful for personal projects, but additional benefits occur in collaborative projects. After several personal assignments and shared discussion forums, we were ready to introduce collaboration. At this point, we would have 36 students adding code to a shared application. As experienced developers may know, adding files to a common code base introduces a variety of common situations including: Added, but missing code files, code conflicts, disappearing edits, and complicated merges.

When code files are completely separate from others, problems are minimal. However, when multiple users edit the same code file—especially on or near the same lines of code, the problems become more challenging.

We introduced both situations by having the students add a top-level menu item for each team. Each team menu item had a sub-menu item for each student taking the user to their own profile page with custom images, css, styles, and appearances. Student-specific pages of course, created minimal conflicts (except sometimes for over-zealous styling), and the shared menu pages created a small, but targeted location to address the more complex issues associated with mutually-edited lively-evolving entities, or *melees*.

This assignment focused on *Creating Views* (and working collaboratively with a large project team).

1. Clone the code from the reference project.
2. Right-click on your projects folder and select *Git Bash Here*. Type the provided URL and hit Enter.
3. In Windows Explorer, right-click on the new solution file to open the project in Visual Studio.
4. When you run the program, more information about this project will show up on the opening page.
5. Add one top-level menu item for each team. Each developer will have their own menu item under their team (with their own menu item, controller action, and view).
6. Add yourself to the menu so users can access your page.
7. Add an action method to the associated controller.
8. Add a view to be returned as your action response.
9. Display any (school-appropriate) content you like on your personal page.
10. See examples under the *Teams* menu option to get started.
11. First, make a small change and then commit your code back into the master branch and push it to the repo. Start with a change to the README.md if you want to be safe.
12. Remember to **add** any new files (e.g., if you included images) before committing your code to your local repository and pushing to the shared master branch.
13. Don't break the build; code can be committed in a partial state but it must compile when you commit.
14. Commit and push small changes. Everyone will be adding their name to the next slot and you could have terrible problems with merging if you don't (a) pull the newest version (b) commit and push your menu changes quickly.
15. Pull new code frequently; the project will be changing quickly as many team members work on a short deadline.

16. Your goal should be to have at least your name added to the menu by the end of our class period tomorrow.

Additional experiences were included throughout the course, but this assignment provided a realistic, or perhaps, even more challenging experience than the typical introduction to Git.

Results from these series of small steps to introduce Git in the native environment for students new to version control are provided in the results section.

5 Evaluating the Introduction of Version Control

Student response to both approaches was measured with a survey. Response to the incremental native introduction of git was given to and completed by 36 students in a graduate-level applied computer science web applications course using ASP.NET 5 (now ASP.NET Core 1.0), MVC 6, and Entity Framework 7. Students varied in programming ability and exposure to version control. Response to the lessons using GitSubmit was given to a smaller class of undergraduate students in a second-year course for new programmers working on their Bachelors of Computer Science. Students were primarily novices in both programming and version control. Five of eleven students completed a preliminary survey; three of these five completed a later version as well.

Students learning Git in a native environment were generally older; all were over 21 and 31% were over 25. 16 of 36 were in their first year of graduate school and 16 were in their second year. 11% were in their third year. Nine participants (25%) were female. Twenty (55%) are first-generation college students.

Students learning git with GitSubmit were generally younger; all were under 25 and three of the five were under 21. Two were juniors and three were seniors. All were male. One was a first-generation college student.

6 Results

Responses to both approaches were positive. On a 6-point Likert scale, students most commonly agreed or strongly agreed with statements such as *the process was easy to use*. For scoring, the following values were used [1].

No answer	1
Strongly disagree	2
Disagree	3
Neutral	4
Agree	5
Strongly agree	6

[1] Given prior experience, we did not anticipate the extent to which the responses would be positive. In later surveys, we may expand the scale to allow more discernment between positive responses.

Student response to the seven general reaction statements below are shown in Figure 6.

1	The process will be applicable to my intended profession.
2	I would recommend this process.
3	I would like to use this process on future assignments.
4	I view this process in a positive light.
5	The process was convenient and an efficient use of time.
6	The process was easy to use.
7	The process distracted from core content and my intended learning.

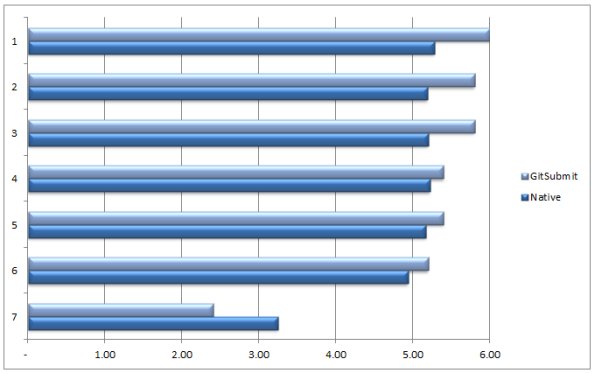


Figure 6: Student response to version control introduction.

Students using GitSubmit to submit their assignments preferred it to the current submission process in our existing LMS and recommended that we propose the process to classes outside computer science. Open-ended comments included:

- *Much more reliable and easier to use compared to [our current LMS].*
- *Keeps track of versions of your work and has the ability to update your work (for revisions and such).*
- *The layout is very logical, especially if you have used version control software before.*
- *It was quick to get all assignments at once.*
- *Easy to grab since it is on my desktop.*
- *I think this program is a fantastic idea, a great solution to the clumsy [current LMS], and I'm sure that it would be useful in many other departments.*
- *It's better overall than [our current LMS].*

Students working with the incremental native lessons provided many comments about things they liked; 35 of 36 students provided open-ended feedback:

- *It was very helpful in collaborating on a project.*
- *Collaboration work can be done easily* [2].
- *Can work on the section of the project individually and collate the work with the team.*
- *Merging and stashing.*
- *Code merging between team members working on different modules.*
- *Easy to commit the code and use it by all.*
- *Easy commands to clone and pull.*
- *You don't have to keep track of all the changes your teammate made in his version.*
- *Easy to clone the project by using GitBash Here.*
- *Branching and merging are easy.*
- *Workflow is flexible.*
- *Good process for learning how to keep a track of your work in the project.*
- *Maintaining the code at one place.*
- *Git was quite easy and it would show all the commands to do for pushing and pulling the code.*
- *Share the code between team members.*
- *Process was easy for initial setup as well as in long run.*
- *It is easy when working on a project. Everyone can work together on the code.*
- *Helps greatly in collaborating to continue the workflow.*
- *Easy to share code and keep teammates updated*
- *Getting acquainted with command line interface.*
- *It is easy to share code and work on it; [whenever] we need we can easily work on code.*
- *It helps in sharing knowledge and also to complete projects.*
- *Fast and progress can be seen.*
- *Easy to use.*
- *I have never tried other tools for the same purpose but the Git looks pretty much easy and efficient.*

28 students offered a comment when asked what they disliked. Sixteen of those students mentioned merge conflicts as their only dislike. Others mentioned initial difficulties, *It was a bit difficult to learn in the start but it made collaboration much easier as I got to know it* and a couple each mentioned difficulties learning commands and synchronizing with the servers.

[2] Ease of collaboration mentioned in 8 responses

7 Conclusions

This paper demonstrates the introduction of version control to novice students in two ways, first through a scaffolded process for those new to both coding and version control and secondly through an incremental process in the native environment for those with some coding experience. The most troubles for native users came from learning commands, understanding the synchronization with the servers, and dealing with merge conflicts, all areas in which GitSubmit provides scaffolding support. Results were positive with students finding version control both very applicable to their profession and, perhaps more surprising for first experiences with Git, easy, useful, and enjoyable to use.

From the initial tests, it appears that GitSubmit may push positive scores to an even higher level. We will continue to develop and extend the integration of Git—via GitSubmit and later via the native environments—into our computer science curriculum.

Further work includes integration and testing of GitSubmit in additional introductory computer science programs at two universities. Concurrently, additional definitions are being added to the associated visual language and the development of additional tools for instructor support for managing and grading large numbers of coding assignments will continue. The progressive lessons will be extended to incorporate additional alternatives such as SourceTree and GitHub Desktop, additional work with native environments including the Git Bash command line, and the further development and testing of GitSubmit. Development of a reusable library of progressive version control assignments will continue, expanding the library to provide set of short, flexible lessons that can be integrated across a variety of existing courses.

References

[1] N. W. Eloe, D. M. Case, and J. L. Leopold, "VeCVL: A Visual Language for Version Control," in *Proceedings of the 2016 International Workshop on Visual Languages and Computing (in conjunction with the 22nd International Conference on Distributed Multimedia Systems (DMS'16))*, 2016.

[2] M. Guzdial, "Software-realized Scaffolding to Facilitate Programming for Science Learning," *Interactive Learning Environments*, vol. 4, no. 1, pp. 001–044, 1994.

[3] D. Wood, J. S. Bruner, and G. Ross, "The Role of Tutoring in Problem Solving," *Journal of Child Psychology and Psychiatry*, vol. 17, no. 2, pp. 89–100, 1976.

[4] S.-K. Chang, "Preface to volume 1," in *Journal of Visual Languages and Sentient Systems*, vol. 1, 2015, p. 4.

[5] S. K. Chang, *Visual languages*. Springer Science & Business Media, 2012.

[6] M. Coccoli, A. Guercio, P. Maresca, and L. Stanganelli, "Smarter universities: A vision for the fast changing digital era," *Journal of Visual Languages & Computing*, vol. 25, no. 6, pp. 1003–1011, 2014.

[7] N. M. Consortium *et al.*, "Nmc horizon report 2014 higher education edition," 2014.

[8] M. Coccoli, P. Maresca, and L. Stanganelli, "Enforcing team cooperation: an example of computer supported collaborative learning in software engineering," in *Proceedings of the Sixteenth International Conference on Distributed Multimedia Systems*, 2010, pp. 189–192.

[9] D. Linville, M. P. Rogers, C. Kelly, C. Spradling, and D. Case, "Profession-based learning through collaboration and vertical alignment with k12, higher education, and industry: panel discussion," *Journal of Computing Sciences in Colleges*, vol. 31, no. 5, pp. 187–189, 2016.

[10] M. Coccoli, L. Stanganelli, and P. Maresca, "Computer supported collaborative learning in software engineering," in *2011 IEEE Global Engineering Education Conference (EDUCON)*. IEEE, 2011, pp. 990–995.

[11] A. Meneely and L. Williams, "On preparing students for distributed software development with a synchronous, collaborative development platform," in *ACM SIGCSE Bulletin*, vol. 41. ACM, 2009, pp. 529–533.

[12] R. Francese, C. Gravino, M. Risi, G. Scanniello, and G. Tortora, "On the Experience of Using Git-hub in the Context of an Academic Course for the Development of Apps for Smart Devices," in *Proceedings of the 21st International Conference on Distributed Multimedia Systems (DMS'15)*, 2015, pp. 292–299.

[13] J. Kelleher, "Employing Git in the Classroom," in *Computer Applications and Information Systems (WCCAIS), 2014 World Congress on*. IEEE, 2014, pp. 1–4.

[14] J. Lawrance, S. Jung, and C. Wiseman, "Git on the cloud in the classroom," in *Proceeding of the 44th ACM technical symposium on Computer science education*. ACM, 2013, pp. 639–644.

[15] V. Isomöttönen and M. Cochez, "Challenges and confusions in learning version control with git," in *Information and Communication Technologies in Education, Research, and Industrial Applications*. Springer, 2014, pp. 178–193.

[16] GitLab, "Code, test, and deploy together with GitLab open source git repo management software," https://about.gitlab.com.

[17] S. Chacon, *Pro Git*, 2nd ed. Berkely, CA, USA: Apress, 2014.

[18] Atlassian, "Bitbucket Cloud Documentation Home," https://confluence.atlassian.com/bitbucket/bitbucket-cloud-documentation-home-221448814.html.

Research on Performance Degradation Assessment Method of Train Rolling Bearings under Incomplete Data

Yong Qin, Xuejun Zhao,Dandan Wang,Zhipeng Wang, Limin Jia
State Key Laboratory of Rail Traffic Control and Safety
Beijing Jiaotong University
Beijing, China

Xuejun Zhao
School of Traffic and Transportation
Beijing Jiaotong University
Beijing, China
15120930@bjtu.edu.cn

***Abstract*— This paper mainly discusses the performance degradation assessment of train rolling bearings under incomplete data, by using the support vector data description (SVDD) and dynamic particle swarm optimization (DPSO).The proposed method is based on the similarity weight for the assessment of the train rolling bearings under incomplete data. Firstly, to obtain effective features of bearing performance degradation from collected vibration data, the local mean decomposition (LMD) is employed to decompose the vibration data. Secondly, the high-dimensionality of features is reduced by the principal component analysis (PCA). And then, on the basis of choosing the kernel parameter and penalty weight, a degradation method based on SVDD is proposed. Finally, the experimental results verified that the proposed method has a better optimization performance than the traditional method and can assess the performance degradation of train rolling bearings under incomplete data.**

Keywords- rolling bearings; performance degradation assessment; support vector; particle swarm optimization.

I. INTRODUCTION

Rolling bearing contributes to a significant portion of train operation. It is a kind of rotating machinery part which is huge in quantity and has complex coupling relationship with other parts. Its working status is directly related to the operation safety of the whole vehicle and has a very important influence on the stable and efficient operation of the train[1]. However, under the harsh circumstance of high speed, heavy load, the working face of the train bearing is very easy to wear, crack or emerge indentation even rupture due to the repeated action of contact stress. According to statistics, only 10%~20% of train rolling bearings can reach the designed life[2]. Therefore, timely and accurate assessment of rolling bearing performance degradation can provide theoretical basis for health management and prediction key components of the train so as to reduce the rate of train failure and promote the active protection of the safety of the train operation.

At present, researches of the train rolling bearing are mainly about the condition monitoring and fault diagnosis. Most of the research results aim to complete the identification and state identification of fault types and the identification results mostly aim at a single time point while ignoring the slight and long-term characteristics of the performance degradation of bearings[3-5] . Therefore, this paper aims at operation data under normal state of the train bearings and discuss performance degradation assessment .

The concept of equipment performance degradation was proposed by Jay Lee[6] in 1970s, mainly including two aspects: feature extraction, performance degradation assessment. At present, in the research of rolling bearings, the above two aspects have been widely studied by many scholars both at home and abroad. In aspects of feature extraction, Fourier analysis, Wigner-Ville distribution, wavelet transform, Hilbert-Huang Transform (HHT) and other time-frequency analysis methods have been proposed[7-10]. In the evaluation of performance degradation there are three main methods based on statistical theory, artificial intelligence and models[11-14]. On the basis of above research results, this paper summarizes the characteristics of the performance degradation of the train bearing. Based on the feature extraction, the performance degradation evaluation method is proposed based on the support vector data description (SVDD) and dynamic particle swarm optimization (DPSO).

II. EXTRACTION BASED ON THELMDAND PCA

In order to provide basic data for the performance degradation assessment of the third section of this paper, the whole feature extraction process of the train rolling bearing is required. Performance degradation assessment focuses on the judgment of the whole process of equipment performance degradation and the description of the dynamic performance of the equipment. Therefore, the feature extraction method needs to satisfy: Keep relatively complete vibration information to facilitate monitoring of the comprehensive state of the whole aspect of the device. The extracted features have high sensitivity, which can reflect the early abnormal state of the equipment. The extracted features can distinguish the degree

DOI reference number: 10.18293/DMS2016-046

of degradation to identify the degraded state. For these reasons, based on the original data of vibration, the vibration data are segmented and local mean decomposition(LMD) is performed firstly. Secondly, based on the decomposition of obtained data, calculating the real-time, statistical and multi-dimensional characteristics. Then use principal component analysis (PCA) to reduce high dimensional feature.

The specific algorithms of LMD and PCA can be referenced in [15] and [16] respectively. When calculating the real time features, to obtain the information of features as much as possible, 12 indexes including RMS, peak, peak factor, kurtosis, RMS amplitude, the average absolute value, skewness, form factor, pulse factor, energy, energy moment, Shanon entropy are calculated. During the calculation of statistical features, data are preprocessed and 4 kinds of common statistical distribution model including the Weibull distribution, gamma distribution, normal distribution and logistic regression are used for distribution fitting and feature vectors are constructed by using the parameters of the distribution model.

III. DEGRADATION ASSESSMENT BASED ON SVDD AND DPSO

This sector mainly discusses the incomplete data of rolling bearings and proposes the performance degradation assessment method based on the SVDD. The data of normal operation are applied for training SVDD hypersphere to find a minimum-volume hypersphere. And the degradation degree of the rolling bearing is obtained by calculating the distance between the feature value of rolling bearing and the hypersphere center. However, owing to the invisibility of multidimensional vector space, the DPSO is applied for the selection of parameters. The flow chart is shown as Fig.1.

A. *Principle of support vector data description*

Support vector data description was originally proposed by Tax and Duin[17]. Given a target object set $x_i \in R^d$ i= 1,. . .,N, the basic idea of SVDD is to find a minimum-volume hypersphere in high dimensional space with center a_F and

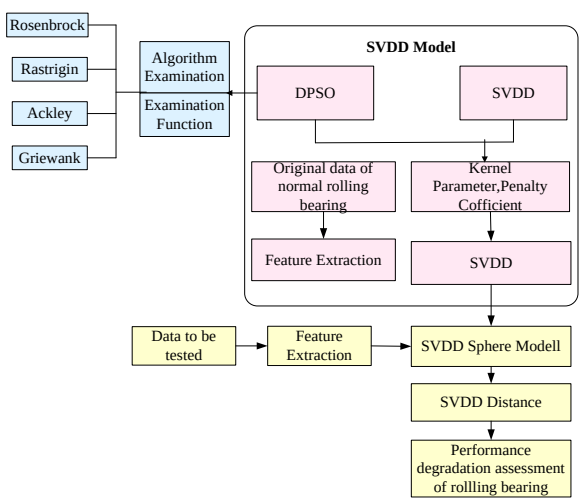


Figure 1 Main steps of bearing performance degradation assessment

radius R to enclose most of the objects, as shown in (1)

$$\text{Minimize } O_p(R, a_F, \xi) = R^2 + c\sum_{i=1}^{N}\xi_i \quad (1)$$

$$\text{Subject to } \|\phi(x_i - a_F)\|^2 \le R^2 + \zeta_i$$

$$\zeta_i \ge 0 \quad \forall \text{ i=1,...N}$$

where c is the penalty weight which gives the trade-off between the volume of the hypersphere and the number of errors. ζ_i are slack variables which allows a probability that some of the training samples can be wrongly classified. ϕ is a nonlinear mapping which maps the input object into a high dimensional feature space F.

The dual problem of (1) is as(2). Where, $K(x_i, x_j)$ is the kernel function.

$$\text{Maximize } O_d(\alpha) = 1 - \sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j K(x_i, x_j) \quad (2)$$

$$\text{Subject to } \sum_{i=1}^{N}\alpha_i \quad 0 \le \alpha_i \le C \text{ i=1,...,N} \quad C \in [1/N,1]$$

In this study, the Gaussian kernel, $K(x_i, x_j) = \phi(x_i)\cdot\phi(x_j)$ $= \exp(-\|x_i - x_j\|^2 / 2\sigma^2)$ is selected. It is because Gaussian kernel has only one free parameter to be turned and is shown to yield tighter boundaries than other kernel choices[18]. Where, α is Lagrange multiplier. According to the Kuhn-Tucker conditions, the objects can be classified into three categories: the object with α_i =0 are inside of the hypersphere; the objects whose $0 < \alpha_i < C$ are on the hypersphere boundary; and the objects whose $\alpha_i = C$ fall outside the hypersphere and have nonzero ξ_i .The objects with $\alpha_i > 0$ are the support vectors. Objectors lying on the hypersphere boundary($0 < \alpha_i < C$) are also called unbounded support vectors. Objects lying outside the hypersphere($\alpha_i = C$) are also called bounded support vectors. The center can be expressed as (3).And its radius R can be determined by utilizing the distance between a_F and any support vector x on the ball boundary (unbounded support vectors), as(4). Finally, for the test object x, the output can be obtained by comparing its distance to the center a_F with radius D in space F. The SVDD decision function is as (5):

$$a_F = \sum_{i=1}^{N_s}\alpha_i\phi(x_i) \quad (3)$$

$$R = (1 - 2\sum_{x_i \in SV_s}\alpha_i K(x_i, x_k) + \sum_{x_i \in SV}\sum_{sx_j \in SVs}\alpha_i\alpha_j K(x_i, x_k))^{\frac{1}{2}} \quad (4)$$

$$D(x) = \left\|\phi(x_i - a_F)\right\|^2 - R^2 = c - 2\sum_{i=1}^{N_s}\alpha_i K(x, x_i) \quad (5)$$

where $c = (1-R^2) + \sum_{i=1}^{N_s}\alpha_i\alpha_j K(x_i, x_j)$ is a constant. For the rolling bearing fault detection, the real-time monitoring data x are accepted as target objects if $D(x) \le D$, which indicates the rolling bearing is normal. Otherwise, it is rejected as an outlier, which indicates the rolling bearing is abnormal. There are two parameters needed to be tuned, C and q. C controls the tradeoff between the volume of the hypersphere and the classification error of the model. It can be tuned to achieve the determined confidence level of the fault detection process. By changing the value of the width parameter $q = 1/2\sigma^2$ in the Gaussian kernel, the description transforms from a solid hypersphere to a Parzen density estimator.

B. *Kernelparameter and penalty weight selection method based on DPSO*

In 1995, James Kennedy and Russell Eherhart proposed the particle swarm optimization (PSO)[19].PSO encourages constructive cooperation and information sharing between particles, which enhance the search for a global optimal solution. However, a potentially dangerous property in PSO still exists: stagnation due to the lack of momentum, which makes it impossible to arrive at the global optimum[20]. To avoid these drawbacks of the basic PSO, some improvements such as the time-varying parameters and random perturbation have been proposed. These improvements can enhance convergence of PSO toward the global optimum, to find the optimum solution efficiently. Therefore, DPSO is proposed to avoid the potential drawbacks. The principal of the DPSO is shown as follows.

Definition: The similarity degrees of two particles have to satisfy the following rules:

(1) $s(i,i) = 1$

(2)When $d(i,j) \to \infty$, $s(i,j) \to 0$

(3)For any particle such as i and j, $s(i,j) \in [0,1]$

Based on the rule, the similarity between i and j is calculated through (6)

$$s(i,j) = \begin{cases} 1, & d(i,j) \le d_{\min} \\ 1 - \left[\dfrac{d(i,j)}{s_{\max}}\right]^{\alpha}, & d_{\min} \le d(i,j) \le d_{\max} \\ 0, & d(i,j) \ge d_{\max} \end{cases} \quad (6)$$

Where $d(i,j)$ is the Euclidian Distance between particle i and particle j. $d_{\max}$ and $d_{\min}$ are constants.

Set iteration times t of $d_{\max}$ and $d_{\min}$ respectively and calculate similarity degree $s(i,j)$ .When the similarity is zero, the inertia weight of the particle is the largest one representing by $\omega_{\max}$ .When the similarity is one, the inertia weight of the particle is the lowest one representing by $\omega_{\min}$ .When the similarity is between zero and one, similarity decreases accordingly. The calculated equations of inertia weight are show as follows:

$$\omega_i = \omega_{\max} - s(i,g)(\omega_{\max} - \omega_{\min}) \quad (7)$$

$$\omega_i = \omega_{\min} - (\omega_i - \omega_{\min}) \cdot \frac{t_{\max} - t}{t_{\max}} \quad (8)$$

Finally, the kernel parameter and penalty weight selection method based on DPSO is made up of 6 steps:

Step1: Generate n locations and initial speeds of kernel parameter or penalty weight

Step2: Evaluate the adaptation of every kernel parameter or penalty weight.

Step3: Confirm the best location p_i of every particle and the global best location p_g

Step4: Calculate particle and global similarity of every penalty or penalty weight according to (6) and calculate particle and global weight according to (7) and (8).

Step5: Update the location and position of kernel parameter or penalty weight.

Step6: If the result satisfies the stopping condition, output the result. Otherwise, turn to Step2.

IV. EXPERIMENTAL RESULT AND ANALYSIS

A. *Original data acquisition*

In this paper, the rolling bearing life vibration data are from the University of Cincinnati intelligent maintenance system (IMS) center. The test apparatus is shown in Fig.2. Four ZA-2115 Rexnord rolling bearings are mounted on the same output shaft with different positions. The speed of the output shaft is 2000rpm. In the shaft and bearing, the radial load of 6000lb is exerted by the spring mechanism. Collect the data by data collection card NI DAQ6062EJ and the sample rate is 20KHz. Collect the data every ten minutes and the collection time is 1 second. The data length is 20480.

This paper takes the original data collecting during every turn of the bearing as the division basis which means that 600 points for a segment. So the data collected in1 second will be divided into 34 segments.

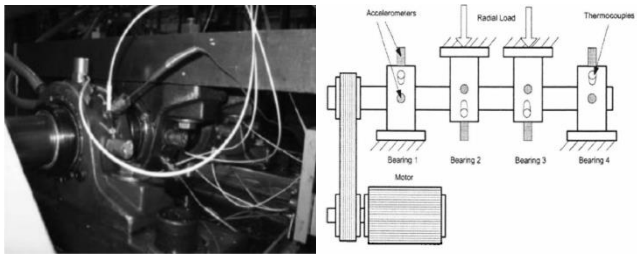


Figure 2. Bearing test equipment and sensors mounting position

B. Features Extraction

*1) Real Time Feature Extraction:*The data of each segment are calculated by the product function (PF) component of LMD and its 12 real time characteristic values are calculated respectively.

Then feature dimension which the 1s original data correspond to is calculated by (9). However, the PF component derived from each segment is different. The minimum decomposition value of the PF component is recorded as the final PF dimension numbers of the (the first 5 PF of each piece of data decomposition result is obtained, the original data per second corresponding to is $34\times5\times12$). To strengthen the difference of the feature value of each stage of degradation, this paper takes the 34 segments of data feature value as the final feature value of the data collected per second(Take the root-mean-square(RMS)value as an example, every group consists of 34 data segment, each segment of data for five PF, each piece of data will get and 5 RMS value corresponding to 5 PF. 34 pieces of data will get 34 RMS corresponding to the first PF, the 34 RMS values are added to obtain the value of RMS the first PF corresponding to. Similarly, other feature value which PF corresponds to can also be obtained . feature dimension in 1s can be figured out.

$$D = N \times M \times L \tag{9}$$

Where D: feature dimensions

N: number of data segments

M: PF numbers of each segment

L: numbers of features

2) Statistical Feature Extraction:Due to the limitation of several statistical distribution for fitting data value is greater than zero, so PF components does not satisfy the requirements obviously, data preprocessing is needed. Because of the good effect of Hilbert transform on earlybearing fault information extraction, do the Hilbert transformation after the decomposition of the PF. Carry out 4 kinds of distributions and fitness testsusing the envelope value after the transformation . This paper makes the distribution fitting test with Minitab software and the value of fit test is calculated. The results show that the Logistic distribution and normal distribution are not good. Therefore, the extracted featureare four including the Weibull shape parameter, Weibull scale parameter, gamma shape parameter gamma scale parameter.

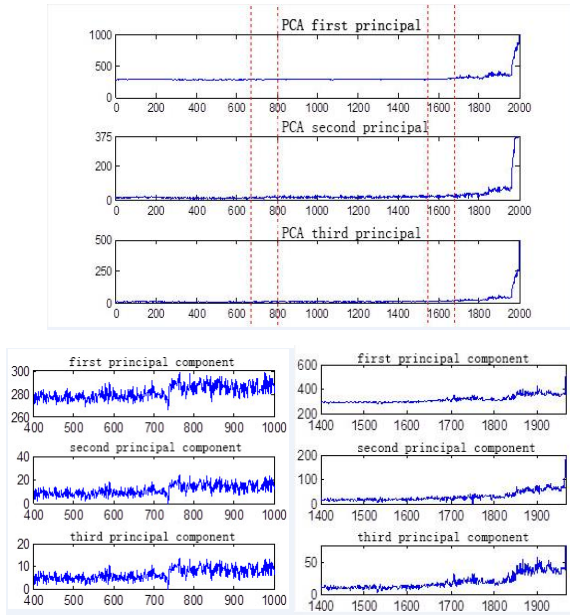


Figure 3. dimension reduction results of PCA (the 1st ~ 3rd principal component)

*3) PCA Dimension Reduction:*As mentioned before, to avoid the trouble of dimensionality and improve the efficiency of the algorithm. On the basis of the extracted multiple real-time and statistical feature, thePCA is used to reduce the dimension.

This paper utilizes the cumulative variance contribution rate method(95%) to choose first 16 feature vectors after dimension reduction. Fig.3 shows the reduced dimension after the first 1~3 main element. It can be seen that the performance degradation trend are well preserved.

C. The Examination of DPSO

To research the optimization ability of the DPSO, several commonly used functions are applied to the performance examination, including Rosenbrock, Rastrigin, Griewank and Ackley. The optimization ability between PSO and DPSO is simulated via four examination functions. The optimization process between the proposed algorithm and the PSO is shown as Fig. 4. Concluding from the Fig.4, both DPSO and PSO converge quickly at the earlier stage. However, Fig.4d) shows that PSO stops declining from the beginning of the iteration, indicating PSO is easily affected by the partial optimum solution. Therefore, the DPSO has a better simulation result.

D. Degradation Assessment

*1) Parameter selection:*Different kernel functions will lead to different high dimensions. Therefore, the performance of SVDD will be affected by the kernel function. As far as the development of the kernel functions, Gauss radial basis function is applied widely in the field of fault diagnosis and its kernel function has characteristic $K(x_i, x_i) = 1$. Therefore, the use of Gauss radial basis function can simplify the objection function and improve the calculation speed. This paper

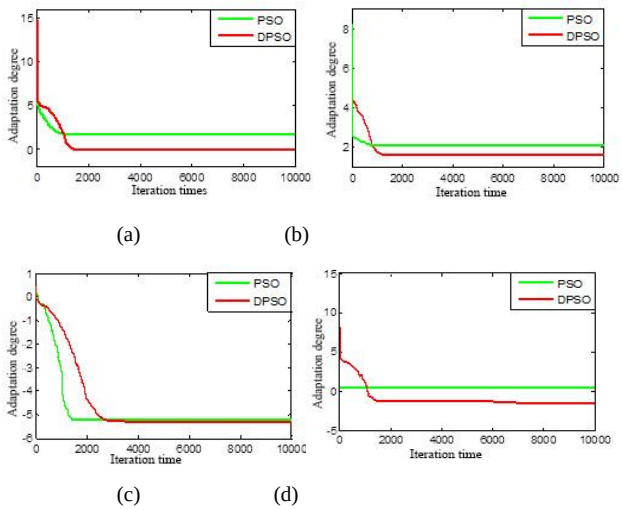


Figure 4. Convergence performance comparison of two methods

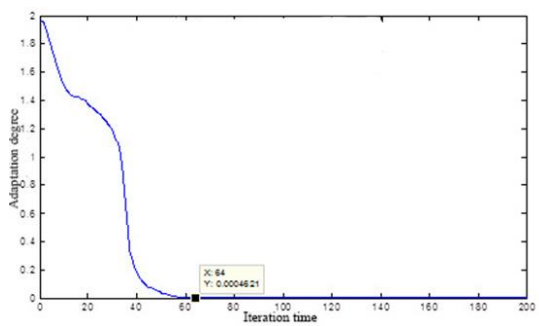
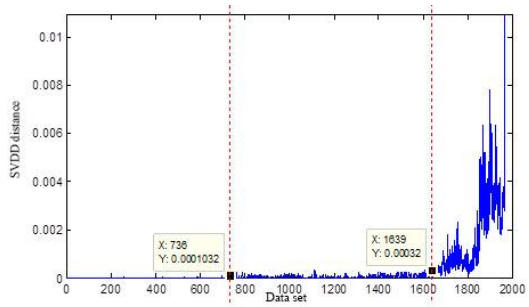


Figure 5. The curve of adaptation

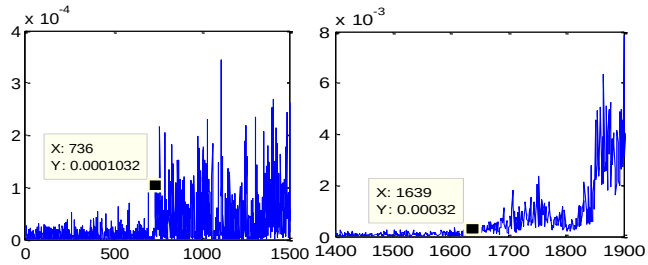
employs DPSO for the selection of kernel parameter and penalty weight and chooses mean square error (MSE) for adaptation and the number of the particle is 20.Besides, the inertia weight is decreasing within the range between 0.95 to 0.4 and the iteration times are 200. The adaptation curve of the SVDD is shown as Fig.5 and the result converges at the sixty-fourth times. The calculated kernel is 15.6832 and the penalty weight is 0.5136.

2) Degradation assessment based on SVDD: The degradation assessment of the rolling bearing is carried out with the feature extracted via PCA. This paper mainly discusses the incomplete data and take the 50 groups normal data out of 2000 groups for training. After standardizing process, the data are input into the SVDD model for finding the center and radius of the SVDD hypersphere. Then data to be tested are input into the existing SVDD model and the degradation degree is obtained by calculating the distance between the data and the hypersphere center.

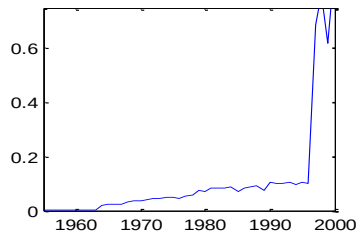
The SVDD distance is shown as Fig.6 and it is divided into three stages. The first stage covers the data between 0 and 735, which means the rolling bearing is normal stage and the SVDD distance is very small and stable. The second stage covers the data between 736 and 1638, which means SVDD distance increases and fluctuates in a larger scale indicating that the rolling bearing enters into an initial stage. The third



(a) SVDD distance



(b) 0-1500 amplification (c) 1400-1900 amplification



(d) 1955-2000 amplification

Figure 6. The curve of performance degradation

stage covers the data between 1639 and 2000 and the SVDD distance increases rapidly up to another scale indicating that the rolling bearing enters into a deep degradation stage.

V. Conclusion

This paper proposes a performance degradation assessment method of train rolling bearings based on SVDD and DPSO. After the decomposition and extraction from the raw vibration signal, with the configuration of the kernel parameter and penalty weight, the SVDD is employed for performance degradation assessment. The experimental results showed that the proposed method is capable to identify the degradation and reflect the health information accurately during the operation of rolling bearings.

Acknowledgment

This research is supported by the National Key Research and Development Program of China (No.2016YFB1200200), the National Key Technology R&D Program of China (NO.2015BAG12B01) as well as the State Key Laboratory of

Rail Traffic Control and Safety of Beijing Jiaotong University (Nos. RCS2016ZQ003, RCS2016ZT016 and RCS2016ZT018).

REFERENCES

[1] Choe H C, Wan Y, Chan A K, Neural pattern identification of railroad wheel-bearing faults from audible acoustic signals: comparison of FFT, CWT, and DWT features[J]. Proceedings of the SPIE - The International Society for Optical Engineering, 1997, 3078:480-96.

[2] Qin T L, Yang Y, Cheng H, Rolling bearing fault diagnosis based on intrinsic mode function energy moment and BP neural network. Journal of Vibration[J], Measurement & Diagnosis, 2008, 28:229-232.

[3] Nabiyev N K, Diagnostics of axle boxes bearings based on identification measuring method[J]. Transactions of University, Karaganda State Technical University, 2010,(1):77-79.

[4] He W, Zhou X, Application of the wavelet-SOFM network in roll bearing defect diagnosis[C]. 2009 WRI Global Congress on Intelligent Systems, 2009:8-12.

[5] He Guangjian, Xing Zongyi, Zuo Cheng, Zhang Yong, Fault diagnosis method for rolling bearing of metro vehicle based on EMD and SVM [J]. Railway Computer Application,2015,(8):1-4,15.

[6] Lee J, Intelligent Maintenance System (IMS) Technologies[OL].[2013-5-30].

[7] Yin Q, Qian S, Feng A, A fast refinement for adaptive Gaussian chirplet decomposition[J]. IEEE transactions on signal processing,2002,50:1298-1306.

[8] Liang Yu, Jia Limin, Cai Guoqiang, Liu Jinzhao, Fault diagnosis method of rolling bearing based on AFD algorithm [J]. China Railway Science,2013,(1):95-100.

[9] Shi Linsuo, Mi Wenpeng, Application of Wigner-Ville-Distribution-based spectral Kurtosis algorithm to fault fiagnosis of rolling bearing [J]. Journal of vibration measurment & diagnosis,2011,31(1):27-31.

[10] Shen Lu, Li Junsheng, Wang Hongjun, Application of improved Hilbert-Huang transform method in gear fault diagnosis[J]. Journal of aerospace power, 2009, 24(8):1899-1903.

[11] Hu T G, Pan Y Z, Zhang J S, Integration of soft and hard classifications using linear spectral mixture model and support vector machines[J]. Spectroscopy and Spectral Analysis, 2011, 31(2):508-511.

[12] Chen Chang, Tang Baoping, L Zhongliang, Degradation trend prediction of rolling bearings based on Weibull distribution and least squares support vector machine [J]. Journal of Vibration and Shock,2014,(20):52-56.

[13] Gao Hongli, Li Dengwan, Xu Mingheng,Intelligent monitoring system for screw life evaluation [J]. Journal of southeast jiaotong university,2010,45(5):685-691

[14] Cong Feiyun, Chen JinDong Guang-ming, Performance degradation assessment by Kolmogorov-Smirnov test and prognosis based on AR model [J]. Journal of Vibration and Shock,2012,31(10):79-82.

[15] Wang Mingda, Zhang Laibin, Liang Wei, Duan Li-xiang. Local mean decomposition method based on B-spline interpolation [J]. Journal of vibration and shock, 2010,29(11): 73-77,254.

[16] Richard A J,Dean W W, Applied Multivariate Statistical Analysis (Sixth Edition) [M]. New Jersey: Prentice Hall, 2007.

[17] Tax D.M.J, Duin R.P.W, Support vector domain description[J].Pattern recognition letters,1999,20(11):1191-1199.

[18] Yang Zhao,Shengwei Wang,Fu Xiao. Pattern recognition-based chillers fault detection method using SupportVector Data Description (SVDD)[J].Applied Energy,2013,112:1041-1048.

[19] J. Kennedy, R. Eberhart, Particle swarm optimization[C]:Neural Networks, 1995. Proceedings., IEEE International Conference on. IEEE, 1995:1942-1948.

[20] Shi Y, Eberhart R. A, Modified particle swarm optimizer[C]. In:IEEE World Congress on Computational Intelligence, Ahchorage, AK, USA, 1998:69-73.

Temporal-Based Risk Forecasting Approach for Key Areas on Surveillance Sensor Networks of High-Speed Railway Transport Hub

Xie Zhengyu
School of Traffic and Transportation
Beijing Jiaotong University
Beijing, China
xiezhengyu@bjtu.edu.cn

Qin Yong
State Key Laboratory of Rail Traffic Control and Safety
Beijing Jiaotong University
Beijing, China
qingyong2146@126.com

Wang Li
School of Traffic and Transportation
Beijing Jiaotong University
Beijing, China
liwang@bjtu.edu.cn

***Abstract*—Key areas risk forecasting plays an important roles in safety management in high-speed railway transport hub. In this paper, a temporal-based risk forecasting approach was considered for key areas on surveillance sensor networks. Computational experiments on a specific key area in high-speed railway transport hub were conducted to illustrate the proposed approach. The results showed the temporal-based forecasting approach is effective and efficient for key areas risk forecasting in high-speed railway transport hub.**

Keywords- Surveillance Sensor Networks; Temporal-Based, Forecasting; High-speed railway transport hub

I. INTRODUCTION

In the last ten years, China high-speed railway had a rapid development. At the end of 2015, a large high-speed railway network has been built with lines in operation amounting to 19,000 km. As an important transferring node, high-speed railway transport hub combines with urban rail transit, taxi, private vehicle, airplane, bus, etc. and has huge passenger distributed every day. Massive passenger flow in high-speed railway transport hub may cause crowd massing situation which bring several potential risks to the safety management of high-speed railway transport hubs.

At present, video surveillances are widely applied in high-speed railway transport hubs for monitoring and detecting passenger status of key areas. Based on the surveillance sensors networks, many approaches for in high-speed railway transport hub safety were proposed in literature.

Several passenger image processing and intelligent detecting algorithms were developed for rapid and accurate detection of passenger flow status [1-4]. To improve detection accuracy in terms of feature extraction, a novel feature of gradient self-similarity (GSS) was present, which was computed from HOG, and was applied to capturing the patterns of pairwise similarities of local gradient patches [1]. In order to accurately and promptly detect potential safety hazard, a modified background model based on Dempster-Shafer theory, and a passenger flow status recognition algorithm based on features of image connected domain were proposed to improve the accuracy and real-time performance of passenger flow detection [2]. For the challenging of abrupt illumination, occlusions, out of field of view, and cluttered backgrounds, a novel tracking framework was developed, which consists of two steps: image shadow removal and tracking by association [3]. A sparse representation based approach is proposed for pedestrian detection from thermal images. The approach first adopted the histogram of sparse code to represent image features and then detect pedestrian with the extracted features in an unimodal and a multimodal framework respectively [4].

Some studies focused on causing and handling of the potential risks [5-7]. The selfish and selfless behaviors were considered as two main factors in evacuation, and selfishness- and selflessness-based model of pedestrian room evacuation was proposed [5]. A multi-grid model was proposed to simulate evacuation with guiders, and the effects of guider type, guider number, guider distribution and guidance strategy on evacuation were discussed [6]. A pedestrian evacuation simulation model based on the extended cellular automata was proposed with the consideration of heterogeneous behavioral tendencies in humans, and the model was applied to optimize buildings to reduce evacuation times [7].

According to literature review above, current studies mainly focused on risks detection and handling. Specific literature on risk forecasting is scarce. In order to accurately detect potential safety hazard hidden in passenger flow, a

DOI reference number: 10.18293/DMS2016-048

hybrid forecasting approach was proposed to forecast the passenger flow status [8]. This approach was only for passenger amount forecasting, and did not consider area risk. Based on the surveillance sensor networks, a passenger flow risk forecasting algorithm based on spatial correlation was developed [9]. The approach was based on the spatial correlation between key area and related areas, and did not consider influence from passenger flow risk temporal change in key area. So in this paper, based on surveillance sensor networks, we forecast passenger flow risk in key area by using itself temporal risk change, and a temporal-based risk forecasting algorithm is developed.

The rest of this paper is organized as follows: The key area risk value calculation is introduced in section 2 and section 3 proposes a temporal-based passenger flow risk forecasting algorithm. Numerical experiments are conducted in section 4 and finally Section 5 covers the conclusion.

II. KEY AREA RISK VALUE CALCULATION IN IN HIGH-SPEED RAILWAY TRANSPORT HUB

Currently, most of high-speed railway transport hubs calculate key area risk by using the passenger flow amount in key area. In the calculation process, three type passenger flow status parameters are adopted. First is current passenger flow amount of key area A_1, which can be real-timely obtained by surveillance sensors in key area. Second is average passenger flow amount of key area in statistic period A_2 and third is max passenger flow amount of key area in statistic period A_3. The last two parameters can be obtained from historical passenger flow amount of key area in statistic period. According to the different relationship among A_1, A_2, A_3, the key area risk value $r(x)$ can be calculated as follows [9]:

(i) While $A_1 \le A_2$, $r(x)$ can be calculated by Eq.(1).

$$r(x) = 3 - 10\left(\frac{A_2 - A_1}{A_2}\right) \tag{1}$$

(ii) While $A_2 < A_1 \le A_3$, $r(x)$ can be calculated by Eq.(2).

$$r(x) = \frac{4A_1 + 3A_3 - 7A_2}{A_3 - A_2} \tag{2}$$

(iii) While $A_1 > A_3$, $r(x)$ can be calculated by Eq.(3).

$$r(x) = 7 + 10\left(\frac{A_1 - A_3}{A_3}\right) \tag{3}$$

III. TEMPORAL-BASED PASSENGER FLOW RISK FORECASTING ALGORITHM

In this section, a temporal forecasting algorithm is proposed to forecast key area risk based on radial basis function neural network in high-speed railway transport hubs.

A. Radial basis function neural network (RBF NN)

RBF NN is a typical feed-forward neural network, which has many merits, such as nonlinear mapping characteristics, self-organized study ability, training fast, and the capability of converging in global optimization and approaching the function in the best way [8]. Simply for its great advantages, RBFNN has been applied in many fields. So in this paper, we propose a temporal forecasting approach based on RBF NN to forecast key area risk in high-speed railway transport hubs.

B. RBF NN design for key area risk forecasting

There are three different layers in the structure of RBF NN, which are input layer, hidden layer and output layer. The structure of RBF NN for temporal-based key area risk forecasting is shown in Fig.1.

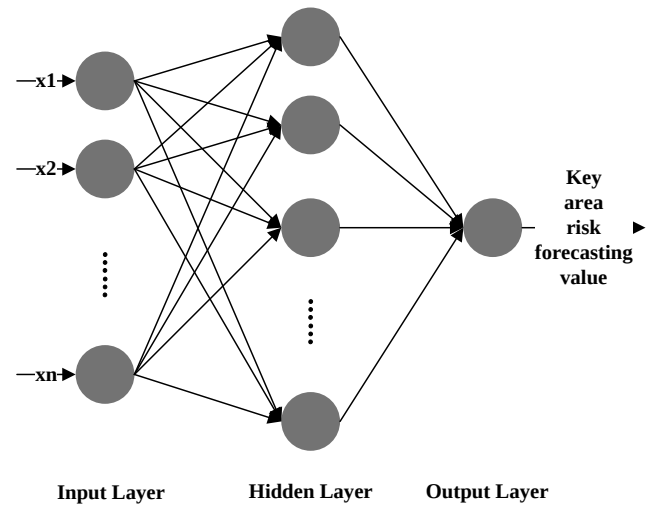


Figure 1. RBF NN structure for temporal-based key area risk forecasting

Main step for RBF NN design is shown as follows:

1) Sample data processing: in order to ensure the input value in the range of 0 and 1, a normalization is need to process the input data. For the time series X, we use Eq.(4) for normalization processing.

$$x_i^{'} = k\left(x_i - \min X\right) / \left(\max X - \min X\right) \tag{4}$$

Where $\min X$ is the minimum value in time series X, $\max X$ is the maximum value in time series X and k is samples normalized scale.

2) Input vector determine: in statistic period, the key area risk values can be regarded as a time series $x = \left\{x_i \middle| x_i \in R, i = 1, 2, ..., L\right\}$, and we can use the risk values before time epoch N to forecast the risk value at time epoch M. Table I is a mapping structure of time series.

TABLE I. MAPPING STRUCTURE OF TIME SERIES

N input value	*M output value*
$x_1, x_2, .., x_N$	$x_{N+1}, x_{N+2}, .., x_{N+M}$
$x_2, x_3, .., x_{N+1}$	$x_{N+2}, x_{N+3}, .., x_{N+M+1}$

N input value	*M output value*
……	……
$x_K, x_{K+1}, .., x_{N+K-1}$	$x_{N+K}, x_{N+K+1}, .., x_{N+M+K+1}$

3) Network parameter: according to a large number of experiments, we finally adopt gauss basis function as the function of hidden layer, k-means clustering algorithm as the center traning method.

IV. NUMERICAL EXPERIMENTS

In this section, numerical experiments are conducted to verify the proposed algorithm. We take a specific key area in high-speed railway transport hub as an example, select time period from 12:00 to 18:00 in a festival day to conduct experiments. We set 10 minutes as a forecasting period, and make a comparison between forecasting risk value and actual risk value. The comparison results are shown in Table II and Fig.2, and the GAPs between forecasting risk value and actual risk value are shown in Fig.3.

TABLE II. COMPARISON BETWEEN FORECASTING VALUE AND ACTUAL VALUE

Hour	Period	Forecasting risk value	Actual risk value	GAP
12:00	1	7	6.8	2.94%
	2	6.9	7	-1.43%
	3	7.4	7.2	2.78%
	4	6.9	7.2	-4.17%
	5	7.3	7.6	-3.95%
	6	7.2	7.4	-2.70%
13:00	7	7.2	6.8	5.88%
	8	7.1	6.8	4.41%
	9	6.7	6.5	3.08%
	10	6.3	6.6	-4.55%
	11	6.7	6.8	-1.47%
	12	7	7	0.00%
14:00	13	7.1	7.3	-2.74%
	14	7.5	7.6	-1.32%
	15	7.9	6	31.67%
	16	8	6.4	25.00%
	17	7.8	6.3	23.81%
	18	7.9	6.7	17.91%
15:00	19	7.7	7.2	6.94%
	20	7.8	7	11.43%
	21	6.9	6.6	4.55%
	22	6.8	6.7	1.49%
	23	7	6.9	1.45%
	24	7.2	7	2.86%
16:00	25	6.9	7.2	-4.17%
	26	7.1	7.4	-4.05%
	27	7.4	7.4	0.00%
	28	7.6	7.8	-2.56%
	29	7.8	7.6	2.63%
	30	7.3	6.9	5.80%
17:00	31	7.5	6.9	8.70%
	32	7	6.6	6.06%
	33	6.8	6.7	1.49%
	34	6.9	6.8	1.47%
	35	7.2	7	2.86%
	36	7	7.3	-4.11%

Note: GAP=(forecasting risk value - actual risk value)/ actual risk value*100%

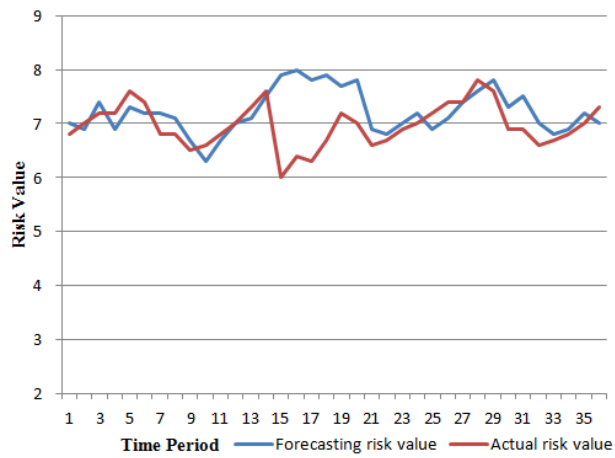


Figure 2. Comparison between forecasting risk value and actual risk value

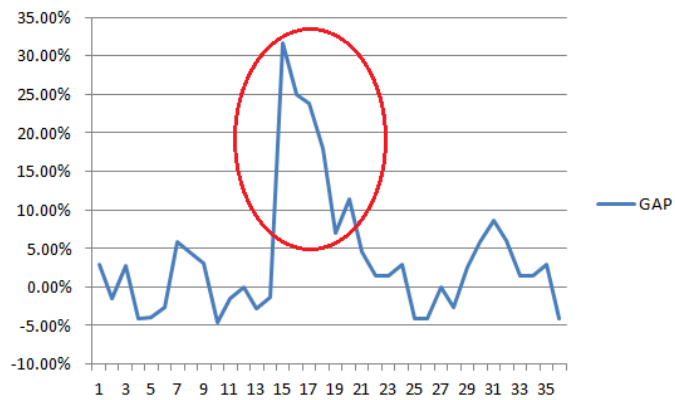


Figure 3. GAP between forecasting risk value and actual risk value

As observed in Table II and Fig.2, the forecasting risk values obtained by our approach are close to actual risk value, except some specific time period. The average gap between forecasting risk value and actual risk value is 3.83%. The variation trend of passenger flow risk is mainly well forecasted by our approach. In the time period 15 to 18, the GAPs forecasting risk value and actual risk value are prominent higher than average GAP (red ellipse area in Fig.3). The forecasting errors result from the passenger flow congestion in related areas of the key area. Our approach is insensitive for congestion situation. Generally, our forecasting approach has a good performance for mass passenger flow in festival days, and can meet the demands of safety management in high-speed railway transport hub.

V. CONCLUSION

In this paper, we considered key area risk forecasting problem in high-speed railway transport hub. Based on the surveillance sensor network of hub, a temporal-based passenger flow risk forecasting algorithm was developed. Numerical experiments on a key area in high-speed railway transport hub showed that the proposed forecasting approach is

effective to forecast key area risk in hub. In future, for our forecasting approach, solving the insensitiveness for passenger flow congestion is an important task for further research.

ACKNOWLEDGMENT

This research was supported by the Fundamental Research Funds for the Central Universities (Grant no. 2015JBM044), the Talented Faculty Funds of Beijing Jiaotong University (Grant no. 2014RC005) and State Key Laboratory Program (Grant no. RCS 2016ZT016)

REFERENCES

[1] Xie Z Y, Jia L M, Qin Y, et al, "Passenger Flow Detection of Video Surveillance: A Case Study of High-Speed Railway Transport Hub in China," Elektronika ir Elektrotechnika, vol.21, pp. 48-53. 2015.

[2] Wu S, Laganière R, Payeur P, "Improving pedestrian detection with selective gradient self-similarity feature," Pattern Recognition, vol.48, pp. 2364-2376, 2015.

[3] Shen X, Sui X, Pan K, et al, "Adaptive pedestrian tracking via patch-based features and spatial–temporal similarity measurement," Pattern Recognition, vol.53, pp. 163-173, 2015.

[4] Qi B, John V, Liu Z, et al, "Pedestrian detection from thermal images: A sparse representation based approach," Infrared Physics & Technology, vol.76, pp. 157-167, 2016.

[5] Song X, Ma L, Ma Y, et al, "Selfishness-and Selflessness-based models of pedestrian room evacuation," Physica A: Statistical Mechanics and its Applications, vol.447, pp. 455-466, 2016.

[6] Cao S, Song W, Lv W, "Modeling pedestrian evacuation with guiders based on a multi-grid model," Physics Letters A, vol. 380, pp. 540-547, 2016.

[7] Li D, Han B, "Behavioral effect on pedestrian evacuation simulation using cellular automata," Safety science, vol.80, pp. 41-55, 2015.

[8] Xie Z Y, Jia L M, Qin Y, et al, "A hybrid temporal-spatio forecasting approach for passenger flow status in chinese high-speed railway transport hub," Discrete Dynamics in Nature and Society, vol.2013, pp. 1-8, 2013.

[9] Xie Z Y, Qin Y, "A Passenger Flow Risk Forecasting Algorithm for High-speed Railway Transport Hub Based on Surveillance Sensor Networks," Journal of Sensors, vol.2016, pp. 1-6, 2016.

Smart Home Control through Unwitting Trigger-Action Programming

Daniela Fogli, Matteo Peroni, Claudia Stefini
Dipartimento di Ingegneria dell'Informazione
Università degli Studi di Brescia
Via Branze 38, Brescia, Italy
daniela.fogli@unibs.it, matteo.peroni89@gmail.com, claudiastefini@hotmail.com

***Abstract*—This paper describes ImAtHome, an iOS application for smart home configuration and management. This application has been built over the framework HomeKit, made available in iOS, for communicating with and controlling home automation accessories. Attention has been put on the design of the interaction with such an application, in order to make the interaction style as much coherent as possible with iOS apps and supporting users without programming skills to unwittingly create event-condition-action rules that, in other similar systems, are usually defined through "if-then" constructs. The results of a user test demonstrate that ImAtHome is easy to use and well accepted by end users of different age and background.**

Smart home; end-user develoment; rule-based programming

I. Introduction

The attention of research scholars and ICT companies is more and more attracted by Internet of Things (IoT) [1] and Ambient Intelligence (AmI) [2], as witnessed by recent conferences, journal special issues and commercial advertisements. These areas involve experts in several disciplines – electronics, artificial intelligence, cloud computing, network infrastructures, and software architectures, just to name a few – who are called on to create and set up a new generation of distributed multimedia systems sometimes referred as "sentient multimedia systems" [3]. However, research on new Human-Computer Interaction (HCI) paradigms is fundamental as well, especially for making AmI environments easy to use and possibly allow their inhabitants, without computer programming knowledge, to install, configure and modify them over time. Therefore, with a particular attention to the smart home, End-User Development (EUD) [4] approaches are being proposed, which are aimed at transforming end users (household members) from passive consumers of sensors, robots and smart devices scattered in the house to active producers of new and possibly coordinated behaviors of such hardware/software components [5].

The idea is indeed to provide the house inhabitants with methods and tools to modify and adapt home behaviors to their needs, in order to cope with the continuous request of user-system co-evolution [3]. This could be achieved by providing users with EUD tools that support them in creating simple commands to be activated manually (e.g. "I am at home, please switch the radio on") or automatically (e.g. "At 7 a.m. rise shutters and activate the coffee machine").

DOI reference number: 10.18293/DMS2016-017

From the analysis of the literature and commercial products [6][7] it emerges the event-condition-action (ECA) rule-based paradigm is the most used in user interfaces devoted to the configuration and adaptation of smart homes by non-expert developers and thus could be considered as a promising solution to create EUD tools in this field. Such tools allow the user to carry out a form of *trigger-action* ("if, then") *programming* [8]. The user is thus guided in setting up the "if" and "then" parts of a rule, by choosing them among lists (filtered-list metaphor), virtual puzzle pieces (jigsaw composition) or components to be put in a network (wired composition) [9].

This paper proposes a new interaction metaphor for rule creation aimed at supporting users to perform trigger-action programming in an "unwitting" manner, that is at helping them create antecedent and consequent parts of the rules, without requiring them to think in terms of "if-then" constructs like computer scientists naturally do. To achieve this goal, the proposed metaphor splits rule creation in two steps, namely the *definition of scenes* followed by the *definition of rules* starting from available scenes. Scenes are sets of device actions that can be also manually activated by the user, thus becoming high-level commands for the house. Rules are defined to make the house able of activating itself some given scenes on the basis of the occurrence of an event, possibly combined with one or more conditions. The metaphor also encompasses a more guided but easier way of defining events and conditions for triggering rules. These ideas have been implemented in an iOS mobile application, called ImAtHome. Terms such as "if", "then" or "do" never appear in ImAtHome, as well as it has been removed the constraint of defining the consequent only after the definition of the antecedent, often present in other similar applications (e.g. IFTTT, Atooma, Tasker, and others).

Another important aspect, often neglected in scientific literature, is the cost of transforming a traditional house into a smart home. Current solutions usually require the interaction with companies that provide global services for smart home installation and maintenance or, alternatively, the acquisition of home automation boxes (e.g., Zipabox, Zibase, Vera, and eeDomus); the latter, in turn, require some "guru" in the family (i.e. a software expert or someone interested in software programming) capable of taking care of system installation and personalization [10].

To overcome this problem, also in this case a smoother approach to home automation is advocated: the idea is to allow

end users to add gradually smart devices to their house, according to emerging needs and economic possibilities. To this aim, we have decided to develop our mobile application over the framework for home automation made available in iOS, namely HomeKit[1]. HomeKit is a framework for communicating with and controlling the smart devices available in a house. It provides the user with a way to automatically discover such devices and configure them. It also makes available functionalities for executing sets of actions to control groups of devices, by possibly triggering them using Siri – the voice-controlled virtual assistant available in iOS.

ImAtHome is thus built on top of HomeKit and proposes itself as a hub application, able to manage all the devices currently available in a smart home, as well as those that will be acquired and included in the future. This paper presents the design and implementation of ImAtHome, as well as a usability study carried out with a group of 14 users of different age and background.

The paper is organized as follows: Section II discusses related works in the IoT and AmI fields, with particular reference to the smart home and user interfaces for their configuration and management. Section III describes the HomeKit framework. Section IV illustrates the development and the operation of the app ImAtHome. Section V provides some implementation details, while Section VI discusses the results of a usability evaluation with 14 users. Finally, Section VII draws some conclusions and proposes hints for future work.

II. Related Works

The idea to enable users to program the behavior of their smart home has been discussed in literature by several research scholars (e.g., [11][12]). For example, in [11], the "Media Cubes" programming language is proposed: it is based on the physical arrangement of infrared remote cubes; they represent abstract functions whose combination leads to the creation of complex behaviors. In the e-Gadgets project [12], instead, a visual editor is proposed, where end users can define "synaptic associations" (cause-effect relationships) between smart appliances available in a home.

More recently, the ECA rule-based paradigm has been proposed in several approaches to EUD applied in IoT or AmI. One of the most complete frameworks for AmI based on a rule-based approach is that described in [13]; a subsequent work of the same authors present three different graphical user interfaces for rule creation [14], even though no usability study is reported. In many cases, the proposed toolkits and languages require users to have some expertise in computer programming and hardware/software technologies. Barricelli and Valtolina have delineated an extension of the ECA paradigm pairing it with the use of formula languages [15]. Coutaz and colleagues [16] have presented a programming environment, called SPOK, which combines rule-based and imperative programming. Demeure et al. [10] described a field study involving 10 households using different home automation systems for a long period of time. In all households, there was always only one member of the family in charge of installing, configuring and managing home behavior modification. This family member was always a male adult, knowledgeable in hardware/software technologies. Therefore, the pure end user, neither expert in software programming nor interested in it, seems practically excluded from the use of such kinds of tools.

Dahl and Svendsen carried out a preliminary comparison among three composition paradigms (filtered lists, wiring composition and jigsaw puzzle composition) for rule creation [9]. From it, filtered lists, where condition-action compositions are obtained by selecting conditions and actions from respective lists, resulted to be the most intuitive for readability; whilst, jigsaw puzzle composition was considered by participants the most playful and engaging type of interaction. The filtered lists metaphor is recently adopted by several commercial and research applications, such as IFTTT, Atooma, Tasker, Locale, and others. Ur and colleagues [8] have proved that IFTTT has a pretty usable interface to create rules, even though these can contain only one event or condition, and only one action. Lucci and Paternò [17] have compared Tasker, Locale and Atooma, all allowing the user to create rules with complex antecedent and consequent parts. In this study, Tasker resulted the best tool in terms of expressiveness and Atooma resulted the easiest to use by end users. The user study reported in [6] compared Atooma and IFTTT in terms of usability and user preferences, by considering both users with a background in computer science and users without this background. The System Usability Scale (SUS) [18] composite score indicated that Atooma has a higher usability; moreover, users appreciated the user interface of Atooma much more than that of IFTTT. A systematic literature review in the IoT and AmI areas is presented in [7], focused on the research works that present tools supporting EUD for smart home configuration and management. From the papers selected through the review, eleven tools have been identified and examined. All tools are based on a rule-based paradigm: end users are supported by visual interfaces to compose events and/or conditions with actions, using structures like 'if-condition(s)-then-action(s)' or 'when-event(s)-then-action(s)'. A qualitative comparison of a subset of these eleven tools is then presented in [7], by considering the design principles for smart home control discussed in [19]. From this analysis, Tasker resulted to be the only tool able to satisfy most of the design principles (six out of seven). However, as also underlined in [17], Tasker appeared as more suitable to users with some knowledge in computer programming than to actual end users. In this paper, we propose a different approach to the creation of ECA rules, which, on the one hand, is aimed to be powerful enough for modeling a huge variety of home behaviors, and on the other hand, would like to support users performing such programming activity in an unwitting manner, as advocated in [20].

Finally, IT companies such as Google, Apple, Samsung and so on, are currently proposing their solutions in this field. However, to allow controlling a variety of devices they may require buying some specific hardware, as in the case of Google Smart Home Media Center, or proprietary accessories as in the case of Samsung. This has consequences on scalability and the possibility for users to create their own rules. On the other hand, Apple proposes HomeKit as a framework for communication with accessories and provides some indications to build apps on it. We have thus chosen to study HomeKit and develop an

[1] https://developer.apple.com/homekit/

application able to exploit it, but open to the interaction with any kind of device compatible with this framework.

III. THE HOMEKIT FRAMEWORK

HomeKit is a framework made available in iOS for communicating with and controlling connected home automation accessories that support Apple's HomeKit Accessory Protocol. Using HomeKit a developer can create complex applications that allow managing the interaction with accessories at a high level, without worrying about low-level technical details. HomeKit is mainly a communication protocol that supports the integration and interoperability of different kinds of accessories.

A. User Interface Guidelines

The user interface of ImAtHome has been developed by following the iOS guidelines for user interfaces and the following more specific HomeKit User Interface Guidelines[2]:

- **Setting up homes** by defining three types of locations: homes, rooms and zones (groups of rooms, such as "upstairs"). Rooms, such as "kitchen" or "bedroom" are the basic organizational concept and will contain the accessories. At least one home must be specified; it will include rooms, and will optionally contain zones. Users must be supported in the creation, modification and deletion of homes, rooms, and zones.
- **Managing users**, who, according to their privileges (Admin or iCloud account holder) may carry out different activities: setting up homes, adding accessories, creating scenes or just adjusting the characteristics of accessories.
- **Adding and removing an accessory** in an easy way, also by means of automatic discovering of accessories. Users must be supported in the configuration of the accessory by assigning it a name, home, room and zone (optional). Users must be able to easily identify the accessory they are configuring.
- **Facilitating the creation of scenes** to adjust the characteristics of multiple accessories simultaneously. Each scene is therefore a set of actions on any number of accessories.
- **Siri integration** to activate scenes with voice commands. HomeKit allows Siri to recognize home, room, and zone names; therefore, Siri can support statements like "Siri, turn off the living room lights".
- **Using a friendly and conversational language**, in order not to intimidate the user with acronyms or technical terms.

Furthermore, HomeKit supports the execution of rules (called "triggers" in HomeKit), which are ways to activate a scene based on conditional relationships concerned with time, location, and the behavior of other accessories.

Therefore, an additional guideline in the HomeKit Developer Guide suggests to **help users set up triggers**, by facilitating as much as possible the creation of the conditional relationships. In the design and development of ImAtHome we have carefully considered this issue.

B. HomeKit Accessory Simulator

HomeKit Accessory Simulator is a tool that allows one to simulate the presence of some accessories in the smart home. Such accessories correspond to those ones that an app would automatically discover in a home. HomeKit Accessory Simulator builds a simulated wireless network to which all accessories added by the developer are connected. Each accessory may have a variety of characteristics to be controlled; through the simulator, the developer can add a characteristic to a class of accessories or a personalized characteristic to a single accessory. Optional characteristics can also be removed. This simulator has been very useful for setting up the experimentation of ImAtHome.

C. Communication in ImAtHome through HomeKit

A variety of companies are developing accessories compatible with HomeKit, such as conditioners, thermostats, light bulbs, cameras, secure locks, carbon monoxide sensors, and so on. As a consequence, several dedicated apps are being developed to control such different devices. Actually, the Apple Store contains at least one app for each accessory class mentioned above and it may also happen that more than one app for the same accessory class are available, usually developed by different producers. The advantages given by the compatibility with HomeKit are that accessories can be controlled through Siri and may be included in the creation of a scene. However, the main drawback is that each accessory keeps on being controlled only by its corresponding app, according to the architecture schematized in Figure 1. Therefore, scenes or rules that involve different types of accessories cannot be created.

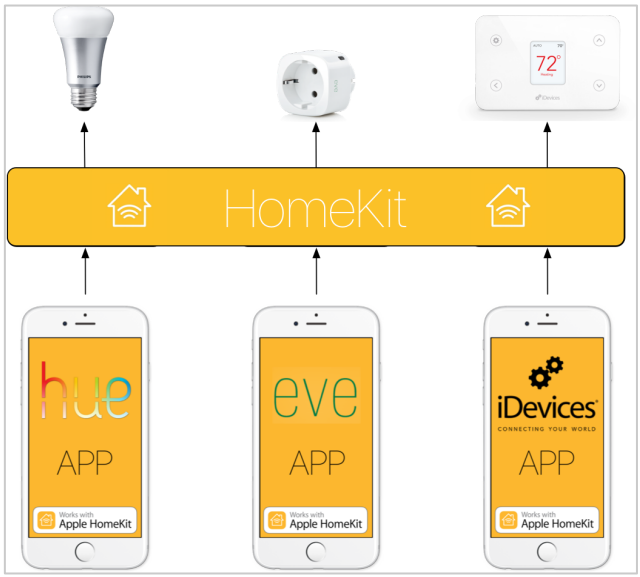


Figure 1. Communication with accessories through HomeKit.

[2]https://developer.apple.com/homekit/ui-guidelines/

The idea underlying ImAtHome is to exploit the common communication protocol for interacting with many kinds of accessories or combinations of them (see Figure 2). In other words, ImAtHome proposes itself as a hub for controlling one's own smart home: all compatible accessories are made available on the smartphone in a unique app, with the same interaction style and the possibility to work in combination one another.

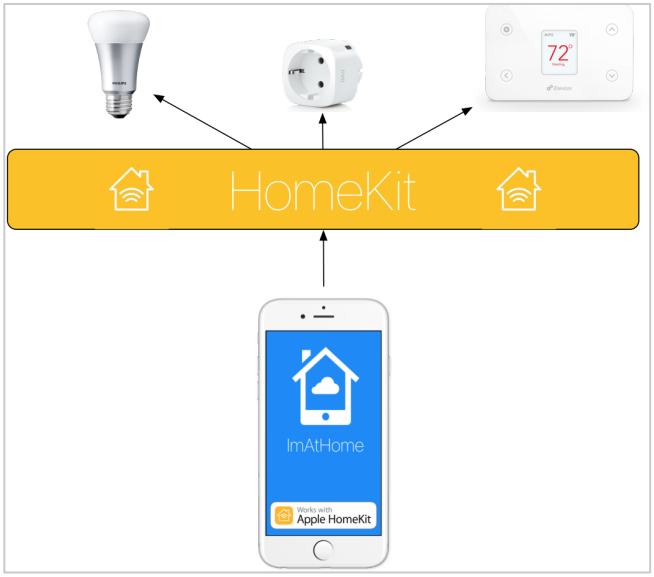


Figure 2. Controlling the smart home through HomeKit with ImAtHome.

IV. The Application ImAtHome

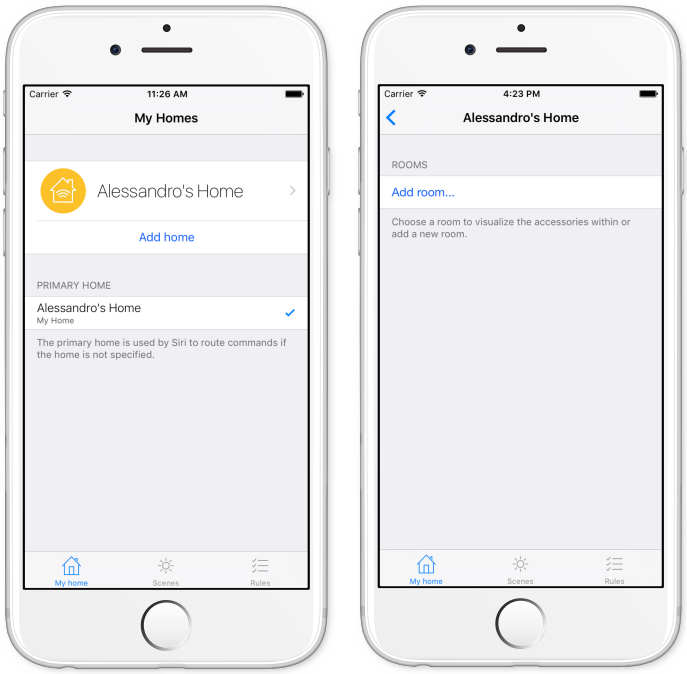
The design and development of ImAtHome has been performed according to a user-centered approach, by involving users in the discussion and evaluation of paper mock-ups and interactive prototypes throughout the software development lifecycle. The following subsections illustrate the interaction with ImAtHome to create homes, scenes and rules.

A. *ImAtHome Interface Structure*

The features made available by HomeKit have guided the ideation of the app structure. First, the interface includes a section, entitled "My home", where the user can access or define a new home (Fig. 3(a)) and its rooms (Fig. 3(b)). By selecting the item "Add room..." in Fig. 3(b), the user can create a room and associate it with a name and an icon among those available (Fig. 4(a)). The new room, "Living room" in the example, is thus added to the list of rooms previously defined by the user (Fig. 4(b)). Accessories are similarly shown as a list associated with a room. The user may bind accessories to rooms inserting their configuration code by means of OCR technology.

Furthermore, the HomeKit database distinguishes between action sets (scenes) and triggers (rules). Action sets may be related to triggers through conditional relationships. Therefore we have decided to add two different sections in the app, one where the user can find or define her/his scenes ("Scenes") and the other devoted to the creation of rules ("Rules"). Scenes are actually sequences of actions that the user may manually activate; whilst rules represent automatic activation of one or more scenes, under some specific conditions (that trigger the rules).

The three sections are accessible through the bottom tab bar.

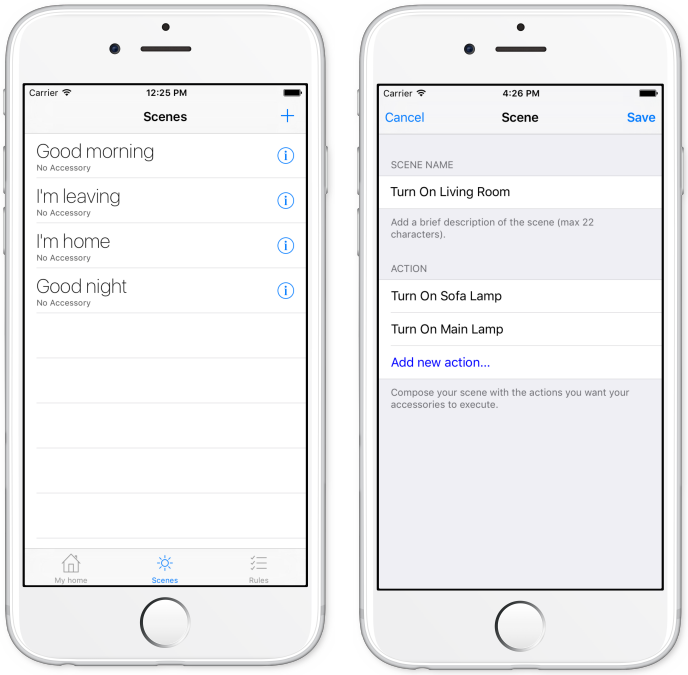


(a) (b)

Figure 3. Section "My home" on the left (a) and screen for adding a room to the home on the right (b).

(a) (b)

Figure 4. Hierarchical navigation in section "My home": screen for room creation on the left (a) and list of rooms on the right (b).



(a) (b)

Figure 5. Section "Scenes" with the list of pre-defined scenes on the left (a) and screen for scene creation on the right (b).

B. *Scene definition and activation*

The section "Scenes" presents a default list of void scenes available in the HomeKit database, that is "Good morning", "I'm leaving", "I'm home" and "Good night" (see Fig. 5(a)). The user may complete them or create a new one by clicking the '+' button in the right upper corner (as it is usually requested in iOS apps to add a new object). In the new screen - Fig. 5(b), the user can give a name to the scene (e.g., "Turn On Living Room" to indicate a scene that turns on the lights in the living room) and define a set of actions (in the example: "Turn On Sofa Lamp" and "Turn On Main Lamp") by setting the characteristics of the accessories involved in each action. When the new scene is saved, the list shown in Fig. 5(a) is updated. Here the user can manually activate the created scene with a tap or by pronouncing its name, thus giving control to Siri for scene execution.

C. *Rule creation*

In ImAtHome, the user may decide to use the pre-defined or user-defined scenes to create rules. Let us suppose that the user would like to create a rule that switches on the lights of the living room (the scene illustrated above) when she arrives at home, but only if it is after 5 P.M. To this aim, she must access the third section of the application and click on the '+' button in the right upper corner. As a consequence, the screenshot in Figure 6(a) is shown. Here the message at the top tells the user to choose among three options to trigger her scenes by "Time", "Position" or "Another accessory". These ones correspond to the three conditional relationships for triggering scenes, which are supported in HomeKit. Selecting one of them allows defining the "event" part of an ECA rule. However, differently from the interaction with other tools, here the user does not need to know that the "if" part of an "if-then" construct must be created. In the example, the user selects the "Position" option. As a consequence, a screen appears where the user

1) defines the details of an event related to her position; in the example in Fig. 6(b) she selects "When I arrive". Then she taps on "Choose a position" and, as a consequence, a map appears centered in the current position of the user; if she is at home, she can simply save that position, otherwise she may look for a specific address through the search bar;

2) defines an additional condition ("after 5:00 P.M." in Fig. 6(c));

3) selects one or more scenes that must be activated. In the example in Fig. 6(d), the user checks "Turn On Living Room".

Note that with steps (2) and (3) the user actually creates the condition part and action part of an ECA rule respectively, without being aware of it. Moreover, differently from other user interfaces for ECA rule definition, ImAtHome requires to define action sets (scenes) first, and then relate them to events and conditions. This allows users to activate scenes manually if needed and use them in several different rules.

V. Implementation

The implementation of ImAtHome has been carried out in Swift, the programming language for iOS presented by Apple in 2014. As already mentioned, it has been built over the HomeKit Application Programming Interface (API), which provides a variety of classes for low-level interaction with home accessories. On activation, ImAtHome creates an object of the class HMHomeManager, made available by HomeKit to add or remove a home to/from its database. For each home, HomeKit creates a database on iCloud, which contains all its objects (rooms and accessories in our case). This database is always synchronized with the user's iOS device; therefore, to show the user the most recent data, the app ImAtHome continuously monitors the database updates. In particular, ImAtHome exploits HomeKit API to 1) discover the accessories available in the environment compatible with the communication protocol, and add them to the database associated to the home; 2) access the properties of the accessories; 3) modify the current values of the accessory properties, thus executing actions (e.g. switch on a light in a given room with a certain lighting level and a specific color). At a higher level, the app is organized in four groups of classes: three main groups manage the behavior and appearance of the three app sections respectively (My Home, Scenes and Rules); whilst, the last group includes all other classes supporting the operation of the main classes.

VI. System Evaluation

A user test has been carried out to evaluate the usability of ImAtHome. The experiment has been conducted by using the iPhone simulator included in the development environment and the HomeKit Accessory Simulator. The Italian version of the app has been used to facilitate the interaction with Italian participants.

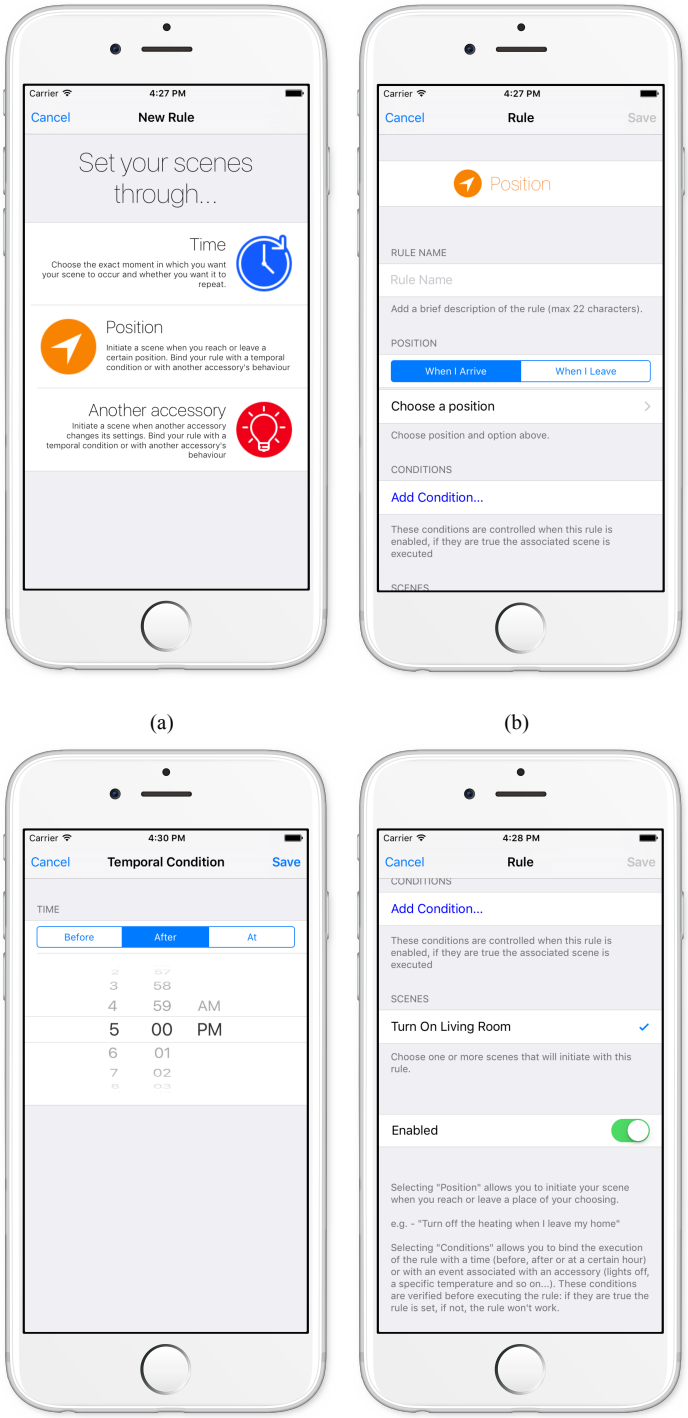


Figure 6. ECA rule creation in ImAtHome.

A. Methodology

The user test involved a total of 14 participants (8 males and 6 females). Their ages ranged as follows: 3 users in 20-24, 8 users in 25-35, and 3 users in 50-65 (average age equal to 33). They held different education degree and represented varied professional backgrounds. They included 8 students, 3 office workers, 1 housewife, 1 unemployed and 1 retired. Twelve participants held an Apple device. As to computer science knowledge, 8 participants declared to have low or medium knowledge, whilst 6 declared themselves as experts. Participants were asked to carry out five tasks of increasing complexity: the first two tasks devoted to the creation of a smart home, with rooms and accessories; the third one for creating a scene; and the last two tasks for defining a simple rule and a complex rule respectively. No previous training was provided to the users.

During task execution, we collected quantitative data related to execution time and percentage of task completion. Since a think aloud protocol was adopted, comments of participants were annotated as well. Then, after the test, we submitted to the participants a post-questionnaire to gather opinions about their experience of use and a SUS questionnaire [18] to obtain an estimation of ImAtHome usability.

B. Quantitative Results

Table I reports the execution times of participants (with average value and standard deviation) and the optimum execution times of one developer. Execution times of participants were always about 3 times the optimum time, except for Task 3 that was the first task devoted to the creation of a scene. The last task was the most complex in terms of actions to perform, but its average execution time was less than that of Task 4, thus demonstrating that ImAtHome is easy to learn. In general, also considering that participants did not receive any previous training on the application, execution times are highly satisfactory. The percentages of task completion reported in the last row of Table I confirm such positive results.

A further analysis on the time spent to carry out the tasks was performed dividing the users in two groups: the former including the 8 participants that declared to have low or medium knowledge in computer technologies, the latter with 6 participants declaring themselves as experts. A t-test was adopted to compare the execution times: no significant difference was found between the two groups, demonstrating once again that ImAtHome allows all kinds of users to easily become “unwitting trigger-action programmers”.

C. Qualitative Results

Comments of the users gathered during test execution highlighted some cosmetic problems that were easily solved after the experiment. Most of the users made positive (and sometimes enthusiastic) observations on the interaction process adopted for creating homes and rooms and populating them with accessories. Some difficulties were encountered in the execution of Task 3, which required to create a scene for switching off all the lights: before choosing the right way, a participant tried at first to modify a pre-defined HomeKit scene; another participant observed that the task was not easy to understand, but he also admitted that he did not read the app instructions first, which instead would have provided him with useful examples; finally, one participant told us that it was not clear when and where to save the scene. Fewer difficulties were encountered in the creation of the first rule (Task 4): only one user had been not able to combine an event (“when I move from home”) with a condition (“if it is after 9.00 P.M.”). As already mentioned, Task 5 was very successful, despite its complexity: indeed, some participants commented that the interaction with the app became familiar after few interactions.

TABLE I. TASK EXECUTION TIME AND COMPLETION PERCENTAGE

User	Tasks				
	Task 1	*Task 2*	*Task 3*	*Task 4*	*Task 5*
U1	02:00	01:48	02:30	03:43	04:54
U2	01:46	02:00	01:30	04:30	02:57
U3	00:40	01:16	01:00	04:30	02:25
U4	01:23	01:30	01:38	03:03	03:35
U5	01:27	01:47	01:55	03:00	02:30
U6	01:20	02:03	02:30	02:15	02:55
U7	00:38	01:03	01:44	01:40	01:53
U8	00:39	01:35	02:00	01:30	02:40
U9	01:23	01:30	01:17	05:34	04:06
U10	00:41	01:30	00:35	02:20	02:10
U11	01:40	01:59	03:10	02:10	03:00
U12	01:27	01:19	01:08	01:45	01:55
U13	01:30	02:47	01:10	04:10	04:12
U14	00:50	01:34	01:29	04:41	02:40
Avg.	**01:15**	**01:42**	**01:41**	**03:12**	**02:59**
SD	00:27	00:25	00:40	01:16	00:52
Opt.	**00:30**	**00:50**	**00:18**	**01:00**	**1:00**
% Compl.	**100%**	**100%**	**93%**	**93%**	**100%**

D. *Findings from post-questionnaire*

The post-questionnaire included the following questions:

1. Did you find the interaction with the app pleasant and funny?
2. If compatible accessories would be present in your home, should you use ImAtHome?
3. Did you find the user interface coherent with the other iOS apps?
4. Do you think that with some limited training ImAtHome would be easier to use?
5. Did you find the language of the application easy to understand?

Answers to the above questions were given on the qualitative scale {"definitely no", "no", "fairly", "yes", "definitely yes"}. We then translated the participants' assignments to the 0-4 numerical scale and computed the average values. The following results were obtained: Q1=**3.07**; Q2=**3.29**; Q3=**3.67**; Q4=**3.29**; Q5=**3.36**. As to Q3 (related to coherence), we gathered the answers of 12 out of 14 participants, since the remaining two participants did not hold any Apple device. Participants that encountered some difficulties in the execution of Tasks 3 and 4 asserted that, with a limited training, the application became very easy to use.

E. *SUS evalutation*

The overall usability of ImAtHome was finally evaluated through the SUS questionnaire, by providing an average cumulative score of the 14 participants equal to 86.6, pretty higher than the conventional threshold equal to 70, adopted for declaring that a system is easy to use [18]. More precisely, Figure 7 shows the SUS scores of all the 14 users; notice that they are all (except one) above or equal the threshold.

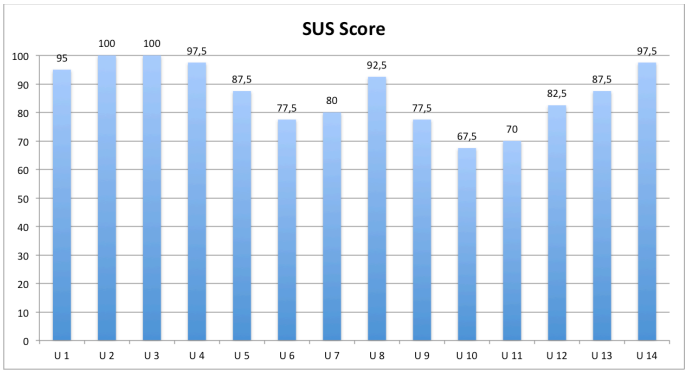


Figure 7. SUS cumulative scores of the 14 users.

VII. DISCUSSION AND CONCLUSION

This paper has presented ImAtHome, an iOS application that allows a home inhabitant, without any programming skills, to control home automation accessories and create scenes and rules for defining complex behaviors of a smart home. The application is scalable, because if a new accessory, compatible with HomeKit, is acquired, it will be automatically recognized by the application and its services will be presented to the users as it already happens for the other accessories. With respect to other approaches proposed in the literature and available commercial tools, ImAtHome allows the user both to manually activate some sets of actions (scenes) or to use them within rules to obtain automatic behaviors of the home. Furthermore, the same scene, once created, can be used several times in different rules; whilst, in other user interfaces based on the "if-then" paradigms, antecedents and consequents of rules must be always explicitly defined.

Integration with HomeKit allows ImAtHome to be used through vocal commands. Indeed, Siri is able to recognize the words associated with accessories and scenes. Therefore, vocal commands can be used to change the characteristics of an accessory and to activate pre-defined or user-defined scenes. An interesting extension could be the creation of a parser for scene and rule creation from the fragments of a vocal command.

As to future work, we are planning to extend the experimentation. Beyond involving a higher number of users, it would be interesting to compare "if-then" or "when-then" interfaces of existing tools (e.g. IFTTT, Atooma, and Tasker) with our "unwitting trigger-action programming" style. Not only measures of user performance, but also users' acceptance and appreciation for the interaction style would represent useful information for future development of these kinds of interfaces.

Another important issue is the extension of the application to the case of multiple user control of a smart home. Indeed, as underlined in [19], more than one person usually inhabits a

home and household activities may be collaborative or in competition (such as TV control or music choice). Therefore, we foresee a future where the collective and participatory evolution of a sentient multimedia system takes place through the simultaneous, but coordinated, intervention of all the interested actors [3]. However, to achieve this goal, some limitations of the current version of HomeKit must be overcome. Currently, it allows associating one's own apple account with a new home, and thus add or modify accessories, rooms, scenes and rules, but it does not allow other Apple accounts to do these activities on the same home. In other words, a home can be shared with other users, but a "guest" user can only control the accessories and activate existing scenes; whilst, she/he cannot actually modify the HomeKit database.

Under the hypothesis that future versions of HomeKit will be released to cope with this issue, we have started to think about the problems that would affect a multi-user approach to smart home control. First, user profiling and permission control should be supported; this would require a usable interface, possibly based on suitable visual languages, like those proposed in [21][22][23]. Second, a variety of social mechanisms, from collaboration to competition, from delegation to reciprocity, should be implemented to stimulate participation. To address this problem, we have proposed the idea of a collaborative application enriched with gamification techniques aimed at motivating all household members to participate in the shaping of their smart home [24][25]. Third, giving household members the possibility to intervene simultaneously in accessory or scene activation and in the creation of rules working on shared spaces may lead to incoherencies and conflicts among rules; suitable solutions must be carefully studied to address these problems.

REFERENCES

[1] L. Atzori, A. Iera, G. Morabito, "The Internet of Things: A survey," Computer Networks, vol. 54(15), 2010, pp. 2787-2805.

[2] F. Sadri, "Ambient intelligence: A survey," ACM Computing Surveys vol. 43(4), 2011, pp. 1-66.

[3] F. Cabitza, D. Fogli, A. Piccinno, "Fostering participation and co-evolution in sentient multimedia systems," Journal of Visual Languages and Computing, vol. 25(6), 2014, pp. 684-694.

[4] H. Lieberman, F. Paternò, V. Wulf, V. (eds.), End User Development. Dordrecht, The Netherlands: Springer, 2006.

[5] F. Cabitza, D. Fogli, R. Lanzilotti, A. Piccinno, "End-User Development in Ambient Intelligence: a User Study," Proc. 11th Biannual Conference on Italian SIGCHI Chapter (CHItaly), ACM, New York, NY, USA, 2015, pp. 146-153.

[6] F. Cabitza, D. Fogli, R. Lanzilotti, A. Piccinno, "Rule-based Tools for the Configuration of Ambient Intelligence Systems: a Comparative User Study," Multimedia Tools And Applications, DOI: 10.1007/s11042-016-3511-2.

[7] D. Fogli, R. Lanzilotti, A. Piccinno, "End-User Development Tools for the Smart Home: A Systematic Literature Review," In: N. Streitz and P. Markopoulos (Eds.): DAPI 2016, LNCS 9749, Springer International Publishing Switzerland, 2016, pp. 1–11.

[8] B. Ur, E. McManus, M. Pak Yong Ho, M.L. Littman, "Practical trigger-action programming in the smart home," In: SIGCHI Conference on Human Factors in Computing Systems, ACM, New York, NY, USA, 2014, pp. 803-812.

[9] Y. Dahl, R.-M. Svendsen, "End-User Composition Interfaces for Smart Environments: A Preliminary Study of Usability Factors," In A. Marcus (Ed.), Design, User Experience, and Usability. Theory, Methods, Tools and Practice, Vol. 6770, Berlin Heidelberg: Springer, 2011, pp. 118-127.

[10] A. Demeure, S. Caffiau, E. Elias, C. Roux, "Building and Using Home Automation Systems: A Field Study," In: Díaz, P., Pipek, V., Ardito, C., Jensen, C., Aedo, I., Boden, A. (eds.) End-User Development. LNCS, vol. 9083, Springer International Publishing, 2015, pp. 125-140.

[11] A. F. Blackwell, "End-user developers at home," Communications of the ACM vol.. 47(9), 2004, pp. 65-66.

[12] I. Mavrommati, A. Kameas, P. Markopoulos, "An editing tool that manages device associations in an in-home environment," Personal and Ubiquitous Computing, vol. 8(3-4), 2004, pp. 255-263.

[13] M. García-Herranz, P. A. Haya, A. Esquivel, G. Montoro, X Alamán, "Easing the Smart Home: Semi-automatic Adaptation in Perceptive Environments," Journal of Universal Computer Science, vol. 14(9), 2008, pp. 1529-1544.

[14] M. García-Herranz, P. A. Haya, X. Alamán, "Towards a Ubiquitous End-User Programming System for Smart Spaces," Journal of Universal Computer Science, vol. 16(12), 2010, pp. 1633—1649.

[15] B. R. Barricelli, S. Valtolina, S., "Designing for End-User Development in the Internet of Things," In: Díaz, P., Pipek, V., Ardito, C., Jensen, C., Aedo, I., Boden, A. (eds.) End-User Development, vol. 9083, Springer International Publishing, 2015, pp. 9-24.

[16] J. Coutaz, A. Demeure, S. Caffiau, J. L. Crowley, "Early lessons from the development of SPOK, an end-user development environment for smart homes," Proceedings of 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication (UbiComp), Seattle, Washington, pp. 895-902, 2014.

[17] G. Lucci, F. Paternò, "Understanding End-User Development of Context-Dependent Applications in Smartphones," In S. Sauer, C. Bogdan, P. Forbrig, R. Bernhaupt & M. Winckler (Eds.), Human-Centered Software Engineering, vol. 8742, Springer Berlin Heidelberg, 2014, pp. 182-198.

[18] S. Borsci, S. Federici, M. Lauriola, M., "On the dimensionality of the System Usability Scale: a test of alternative measurement models," Cognitive Processing, vol. 10(3), 2009, pp. 193-197.

[19] S. Davidoff, M. K. Lee, C. Yiu, J. Zimmerman, A. K. Dey, "Principles of Smart Home Control," In: Dourish, P., Friday, A. (eds.) UbiComp 2006: Ubiquitous Computing. LNCS, vol. 4206, Springer, Berlin Heidelberg 2006, pp. 19-34.

[20] M. F. Costabile, P. Mussio, L. Parasiliti Provenza, A. Piccinno, "End users as unwitting software developers," In Proceedings of the 4th international workshop on End-user software engineering (WEUSE '08). ACM, New York, NY, USA, 2008, pp. 6-10.

[21] M. Giordano, V. Loia, G. Polese, G. Tortora, "A system for user friendly pervasive computing management," In Proceedings of the 3rd International Conference on Intelligent Environments (IE'07), Ulm, Germany, September 2007, pp. 282-287.

[22] M. Giordano, G. Polese, G. Scanniello, G. Tortora, "A system for visual role-based policy modelling," Journal of Visual Languages & Computing, vo. 21(1), 2010, pp. 41-64.

[23] L.Caruccio, V. Deufemia, C. D'Souza, A. Ginige, G. Polese, "A Tool Supporting End-User Development of Access Control in Web Applications," International Journal of Software Engineering and Knowledge Engineering, vol. 25(2), 2015, pp. 307-331.

[24] F. Benzi, F. Cabitza, D. Fogli, R. Lanzilotti, A. Piccinno, "Gamification Techniques for Rule Management in Ambient Intelligence," In: B. De Ruyter, A. Kameas, P. Chatzimisios and I. Mavrommati (Eds.), Ambient Intelligence, Springer International Publishing, 2015, pp. 353-356.

[25] D. Fogli, R. Lanzilotti, A. Piccinno, P. Tosi, "AmI@Home: a Game-Based Collaborative System for Smart Home Configuration," In: Proceedings of International Working Conference on Advanced Visual Interfaces (AVI '16), ACM, New York, NY, USA, 2016, pp. 308-309.

AP Positioning for Estimating People Flow as Origin Destination Matrix for Smart Cities

Pierfrancesco Bellini, Daniele Cenni, Paolo Nesi
Distributed Systems and Internet Technology Lab, http://www.disit.dinfo.unifi.it ,
Department of Information Engineering, http://www.dinfo.unifi.it
University of Florence, http://www.unifi.it, Florence,Italy phone +39-3355668674
Email: {pierfrancesco.bellini, daniele.cenni, paolo.nesi}@unifi.it

Abstract— Controlling and regulating people flows and the access to the city services are major topics in the context of Smart City management. Flow surveillance provides valuable information about city conditions, useful for not only monitoring and controlling the environmental conditions, but to optimize the exploitation of various city services. In this context, it is mandatory to develop tools for assessing people flow. This paper presents a methodology for an effective placement of counter sensors, to model flows with a statistically significant precision rate. Comparative analyses are conducted with respect to real data (i.e., cab traces) of the city of San Francisco. Several different placing methodologies of Wi-Fi access points have been tested and compared, to minimize the cost of AP installation. The research work described in this paper has been conducted in the scope of the EC Horizon 2020 funded project Resolute (http://www.resolute-eu.org) and for Sii-Mobility.
Keywords: people flows; smart city; WiFi Access Point location

I. Introduction

The optimization of services for the citizens is one of the most challenging activities of the Smart Cities. City services can be related to mobility, government, energy, cultural events, commercial, environment, etc. Among the services, mobility is a commodity; thus, transportation and mobility analyses are valuable aspects always considered for an effective definition of Smart City. According to [Giffinger et al. , 2007], Smart Mobility is among the key factors of a modern Smart City, including local and international accessibility, availability of ICT infrastructures, sustainable, innovative and safe transport systems. [Caragliu et al., 2009] include traditional transport communication infrastructures among the essential requirements for Smart Cities.

In the context of mobility, traffic/flow analysis is a major prerequisite for planning traffic routing. Thus, it is a central part of the so called Intelligent Transportation Systems (ITS) for managing public transportation. Traffic flow analysis is commonly used to ease the transportation management, for regulating the access control to the cities, for Smart Parking, for traffic surveillance providing information about road conditions and travel, or for monitoring and controlling the environmental conditions, such as harmful emissions (e.g., CO2, PM10, ozone). The European Commission indicates, among the main topics that should be considered with special attention in the framework of the CARS 2020 process, the implementation and promotions of ITS, including Smart Mobility [CARS 2020]. Some of the techniques adopted for traffic monitoring and management can be declined for people flow analysis and support in the city. It is very important to know the movement of people within a certain precision, and detecting where and how people are crossing the city and using its services by using different kind of moving solutions: car, bike, walking, taxi, car sharing, buses, tram, etc., targeting services into the city [Bellini at al., 2014]. Typically, the telecom operators are not capable to provide this kind of information. They know the number of people connected to each cellular cell at a given time slot during the day, and not how people move in city.

At this regard, specific tracking services for mobile phone IDs are needed and, when applied, the citizens have to be informed via an informed consent (e.g.,

DOI reference number: 10.18293/DMS2016-021

terms of use, policy privacy). In order to measure/derive the typical people flow, the estimation of the so called OD matrix (Origin Destination Matrix) could be needed. The OD matrix presents on both axes the city zones, while the single element may contain the number of people (or the probability) of passing from the zone of origin and reaching the zone of destination, in the day or in a given time window. The OD matrix estimation via observations is very relevant for traffic flow prediction and management, in particular for (i) planning optimized routes predicting shortest and viable paths, (ii) providing info-traffic services on desktop or mobile devices, via the so called Advanced Traffic Management Systems (ATMS).

Most of the OD matrixes can be time dependent, and thus their dynamic real-time estimation may be needed, or at least the estimation of their values along the day and week, (e.g., typical sampling period is every 15 minutes). Moreover, real-time OD matrixes can be continuously estimated during the day or in a temporal window. On the other hand, OD matrixes can be sensitive to traffic conditions. Their values are of primary interest if they represent the maximum or at least sustainable values, disregarding values when the traffic system cannot sustain the traffic. Thus, pre-calculated OD matrixes can be used as default descriptors of the traffic conditions for plan estimation.

These solutions are called Advanced Traveller Information Systems (ATIS). In practice, we would like to measure the typical people flows from the several zones of the city. In the context of traffic flow, some methods make use of parametric estimation techniques (e.g., Maximum Likelihood, Generalized Least Squares, Bayesian inference). ML methods minimize the likelihood of observing the OD matrix and the traffic counts. Other methods based on traffic count include Combined Distribution and Assignment (CDA) [Cascetta et al., 2001], Bi-level Programming [Doblas et al., 2005], [Kim et al., 2001], Heuristic Bi-level Programming [Lundgren et al., 2008], Path Flow Estimation (PFE) [Nie et al., 2005], or Neural Networks [Gong, 1998]. For example, [Ashok and Ben-Akiva, 2000] used a Kalman filtering technique to update the OD matrix. Time dependent offline estimation deals with time-series of traffic counts. Typically, building an OD matrix for mobility requires installing devices to count the single vehicle in the traffic (and eventually record the speed of each vehicle) on the road. A traffic counter is a device that records vehicular data (i.e., speed, type or weight). At this regard, the US Federal Highway Administration defines three main traffic counting methods: human observation (manual), portable traffic recording devices and permanent automatic traffic recorders (ATR). Thus, at level of traffic flow observation several different techniques are used: video cameras, pneumatic road tubes, piezo-electric sensors embedded in the roadway as inductive loop detectors, magnetic sensors and detectors, microwave radar sensors, Doppler, passive infrared sensors, passive acoustic array sensors, ultrasonic sensors, laser radar sensors. Most of these sensors use intrusive technologies and require pavement cut; in some cases lane closure is required, the devices are sensitive to environmental conditions and require an expensive periodic maintenance.

Some of the above mentioned techniques can be used to produce vehicle classification (e.g., rural cars, business day trucks, through trucks, urban cars). Recently, other techniques have been adopted as RFID, Bluetooth, Real Time Location System (RTLS) and Wi-Fi access points [Danalet et al., 2012], [Patil et al., 2015]. In some cases, the position of vehicle can be monitored from the GPS position of mobile devices installed on the vehicle, or simply by using smartphone navigators (e.g., Google Maps, TomTom, Waze), thus providing crowd sources positions and velocity of the vehicles. In these two cases, the track of position is agreed with the users that install the device or run the mobile application on the smartphone or navigator. RFID is quite unsuitable to detect devices because of the small range of action. Bluetooth can be more suitable but expensive, since specific station to collect the passage is needed. Wi-Fi access points are less reliable in detecting the presence of motorized sources with respect to physical devices, and GPS methods.

We decided to use Wi-Fi Access Points as counter devices: given the high distribution of mobile devices and the low cost of a Wi-Fi AP, and the fact that a large number of APs is already installed in the city. This solution is quite cheap and easy to implement, also considering that many municipalities offer free Wi-Fi connectivity, and the needed coverage can be easily obtained. Therefore, the identification of the best placement of Wi-Fi Access Points, as detectors

for measuring the OD matrix is very important. The positioning of APs aims at limiting the costs by obtaining acceptable measures for the OD matrix. That means to obtain an OD matrix with good approximations with respect to a massive and systematic measuring of the whole city flow by humans and/or specific sensors. This paper presents a study and methodology for the positioning of Wi-Fi APs according to the zone adopted for the OD matrix needed estimation. The proposed study and solution has been validated by exploiting the data set introduced in [Piorkowski et al., 2009] which covers cab mobility traces, collected in May 2008 in San Francisco.

The proposed approach assesses the present distribution of Wi-Fi Aps on San Francisco to collect flows data with mobile device discovering. On the other hand, the proposed method could be used to suggest better positioning and/or completing the positioning of the city Wi-Fi APs distribution to be used also as sensors, for the OD matrix estimation. The research work described in this paper has been conducted for RESOLUTE EC Horizon 2020 project (http://www.resolute-eu.org).

This paper is structured as follows. In Section II, the definition of the OD matrix is reported together with the analysis of data collected in San Francisco. Section III presents the proposed approaches for positioning the APs and their comparison. In Section IV, results are analysed. Conclusions are drawn in Section V.

II. Origin Destination Matrix and Reference Data

According to the introduction, the main goal of the presented research has been to identify and validate a method and tool for the positioning of Wi-Fi APs, with the aim of using them as sensors/observers for estimating the OD matrix. In this context, the OD matrix representing flows among the zones of the city, considered as zip codes z is defined as

$$OD_{n,n} = \begin{pmatrix} z_{1,1} & \cdots & z_{1,n} \\ \vdots & \ddots & \vdots \\ z_{n,1} & \cdots & z_{n,n} \end{pmatrix} \qquad (1)$$

where $z_{i,j}$ represents the total number of traffic counts from z_i to z_j (i.e., how many cabs moved from z_i to z_j) defined as

$$z_{i,j} = \sum_{t \in T} n_t(i,j) \qquad (2)$$

and T is the set of unique cab traces, $n_t(i, j)$ is the number of traffic counts from z_i to z_j for trace t.

The validation of any AP positioning in the city for the people count (or traffic flow) is not a trivial task. In principle, one should be capable to install the APs in the city in certain positions and demonstrate, by measuring on the real context, that their position and data collected is strongly correlated with the effective number of people and thus of flows among the different areas of the city. This approach is very expensive and unfeasible for a number of configurations. Therefore, we adopted an indirect method described as follows.

The data set introduced in [Piorkowski et al., 2009] includes the trace flows of cabs in the area of San Francisco, collected in May 2008. The dataset reports all the cab traces in that area, providing fine GPS positioning of them. In **Figure 1**, the trace flows are reported on the city map (for a particular day in the time range 8:00 am - 9:00 am). The data set consists of 446,079 traces based on about 11.2 millions of single GPS points collected by cab movement, not only in the downtown of San Francisco, but spanning on the whole city neighbourhoods. On the other hand, the areas at higher density are those in the downtown, coherently with what one could have from the movements of pedestrians in the city. This area can be identified with about 13 central zip code areas.

Figure 1. Trace flows in San Francisco on a working day of May, 8:00 AM - 9:00 AM

In order to perform an effective data analysis and visualization, some web tools for viewing and comparing flows in different scenarios were developed. At this regard, OD matrix and thus flows

among zip areas are represented with a chord diagram to put in evidence single and aggregate contributions to the total flow count among the various city zones (in **Figure 2,** the chord diagram is reported for the central part of the city with 13 zip code areas). An interactive version of this tool is accessible at http://www.disit.org/6694.

On the user interface, the user can select a time interval on the day to visualize its related chord diagram, which is constituted by circular sectors, each of them representing a city area; passing the mouse over a sector provides additional information about the traffic counts originated from it towards other zip areas. In this manner, it is possible to depict in a compact and intuitive way the traffic flows among the various zones. Moreover, for San Francisco we know the structure of the city and the position of the APs in the downtown (see **Figure 3).** The positions have been taken from OpenWiSpots, http://www.openwispots.com , with GPS positioning. They consist of 494 Wi-Fi Aps offered by city services, from a total of 983 APs, (e.g., coffee shops, hotels, restaurants, libraries, bars, bookstores, grocery stores). Therefore, we supposed to use the Wi-Fi network of San Francisco to estimate the passage in the city by the mobile phones according to their MAC address.

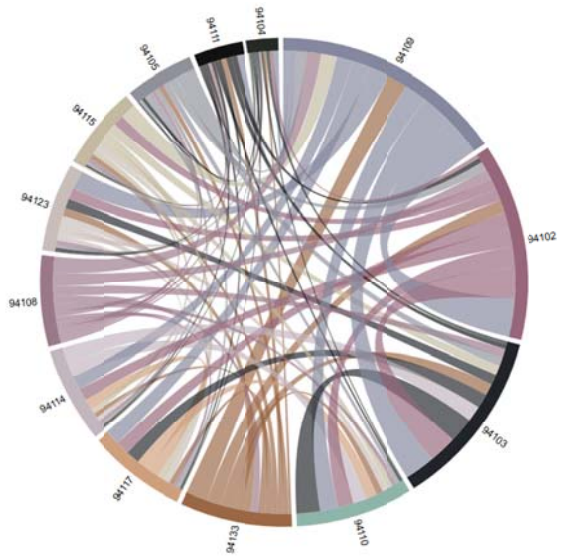


Figure 2. San Francisco OD matrix as chord among the 13 central ZIP areas of the city (real cab flows)

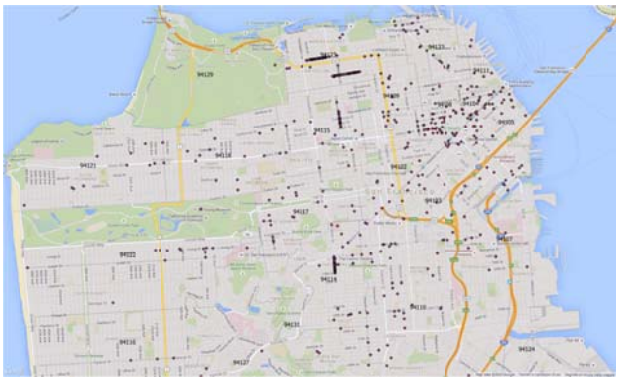


Figure 3. Distribution of real Wi-Fi APs in San Francisco

This solution can be implemented by collecting the first discovery by mobile of the Wi-Fi AP that can stream them as log data to a central server that also anonymizes the MAC. In alternative, some of the APs or aggregators of APs may compute the anonymization algorithm, based on a hash code of the identifiers. Once detected the passages of flow on the APs, the OD matrix can be computed. As a first approximation, we assumed to have the possibility of detecting the flows by using the present APs, by capturing the real traces passing within a distance of 25 meters from the AP position. The proposed approach is a sort of partial simulation based on real data about traffic flow, that is more realistic than producing fully simulated data. It is obvious that the real data captured by the APs would be probably only a part of the real traffic of people passing close to them. On the other hand, it is reasonable to have the simulated measures performed as strongly correlated to the real effective numbers.

As a general consideration, only 1,470,091 points were found to intersect with the real APs positions, which downtown are 1,418,207 with respect to 494 APs. Therefore, in this manner, we assessed the available distribution of Wi-Fi APs in San Francisco, in order to collect people flows data with their mobile device. Once obtained the observations by intersection of traces with respect to the APs areas, an estimated OD matrix on the basis of the APs has been obtained as reported by the chord diagram in Figure 4a. In Figure 4b, the matrix of difference from the OD matrix of Figure 2 and that of Figure 4a is reported, the differences from back abnd forward are not perceivable.

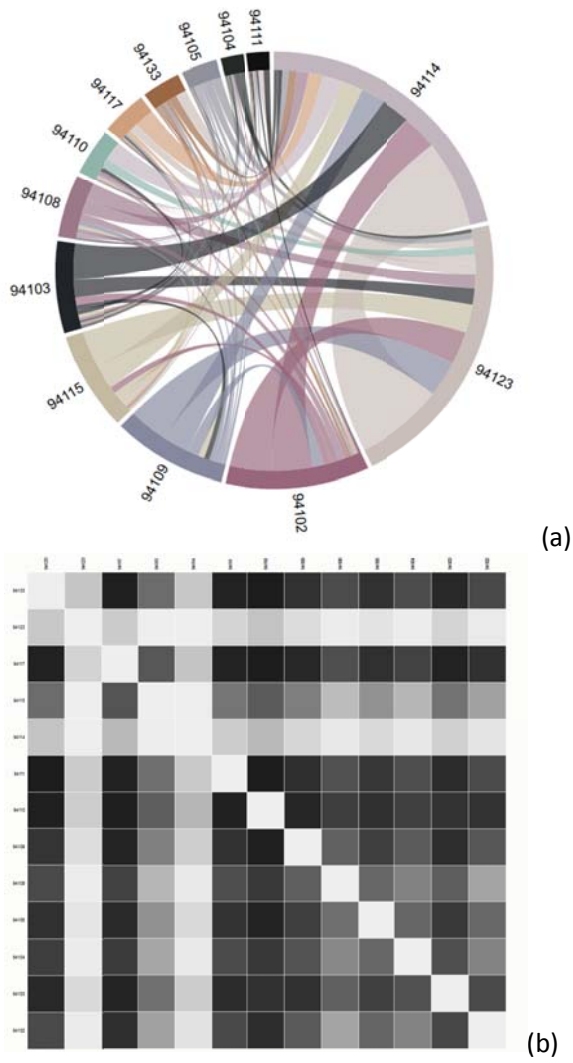


Figure 4. (a) Chord diagram of flow counts with real Wi-Fi APs in the city centre; (b) Difference matrix among OD matrices of real flows and estimated with real Wi-Fi APs in the city centre

The difference matrix compares the real trafc flow data with respect to the flow that would be estimated by using the present APs distribution in the city. The difference values are reported with a grayscale (the higher the difference, the higher the darkness of the matrix element). The two distributions of Origin-Destination are uncorrelated (a correlation of 0.12 has been measured, see Table I). This result demonstrates the unsuitability of the present distribution of the APs for collecting flows. On the other hand, their positioning was not performed with the aim of measuring and observing the flows.

III. Adopting AP Positioning Models

A different positioning of the APs in the zip areas could create better correlation and thus better precision for the estimation of flows in the city. To this purpose, a number of different methods for AP positioning and thus for flow observations have been adopted and tested, taking them from the literature of the classical traffic flow observations strategies by humans. We started by creating a uniformly distribution grid of APs, placed at the middle of each street. In all cases, each AP was considered as a circular buffer with 50 m of diameter. The resulting AP set, consisting of 14,959 APs (a number of devices that is surely too high to be affordable), was further reduced using different strategies as reported in the following. Moreover, the reduction is also reasonable since a uniform distribution in all the zones of the city is not feasible. There are many zones in which the flows are very low, at least in the simulation data taken into account. On the other hand, the positioning of the APs in low flow areas is not efficient.

Also, a flow prediction strategy should be able to tell where to place traffic sensors, and how many sensors to use, providing a tuning strategy for selecting the required set of sensors, with the aim to minimize the number of traffic sensors and the costs of periodic maintenance of the monitoring infrastructure. In this section, we provide some alternative strategies of AP placement. The possible cases for distribution are the following.

a) *Random APs*: identification of the streets with the highest trace flow rate (those that have at least 30,000 traces) and then randomly selecting 400 APs from the AP grid described above (see Figure 6a for the OD matrix). This set of APs is a subset of the set selected in case (b).
b) *High Traffic APs*: identification of the streets with the highest trace flow rate (those that have at least 3000 traces) and then selecting all the APs intersecting those traces, thus resulting in 804 APs (see Figure 6b for the OD matrix).
c) *High Traffic APs (zip boundaries)*: identification of the streets with the highest trace flow rate (those that have at least 3000 traces) and then starting from the 804 APs of case (b) and considering only those within 300 m of the zip boundaries, thus resulting in 448 APs (see Figure 6c for the OD matrix). This set of APs is a subset of the set

Model		Coefficient	Std. Error	t-statistic	p-value	Correlation	# APs
Real APs	β	280393.858	19874.972	14.108	0.000	0.446	983
	α	9.448	0.543	17.400	0.000		
Real APs (cc)	β	1598664.580	116546.825	13.717	0.000	0.12	494
	α	1.714	1.141	1.502	0.135		
(a) Random APs (cc)	β	690144.338	75267.849	9.169	0.000	0.835	400
	α	52.921	2.813	18.816	0.000		
(b) High Traffic APs (cc)	β	684144.945	52950.289	12.921	0.000	0.915	804
	α	10.942	0.389	28.114	0.000		
(c) High Traffic APs (bn, cc)	β	1101641.803	86354.599	12.757	0.000	0.687	448
	α	13.586	1.159	11.727	0.000		
(d) High Traffic APs 400 (cc)	β	810743.094	70801.471	11.451	0.000	0.835	400
	α	24.429	1.297	18.829	0.000		
(e) High Traffic APs (bn, cc)	β	748987.390	58260.615	12.856	0.000	0.892	400
	α	39.960	1.634	24.453	0.000		

Table I -- COEFFICIENTS — REAL APS, CC = CITY CENTER, BN = ZIP BOUNDARIES (WITHIN 300 METERS)

selected in case (b).

d) *High Traffic APs (top 400)*: identification of the streets with the highest trace flow rate (those that have at least 3000 traces) and then starting from the 804 APs of case (b) and considering only the top 400 APs (see Figure 6c for the OD matrix). This set of APs is a subset of the set selected in Case b.

e) *Real augmented with selected High traffic APs*: the real distribution of the AP in San Francisco's downtown was integrated with the top 300 AP from case (d) with the highest traffic rate. This set was then cleaned up by removing those APs that were found to be at a distance less or equal than 50 m from the real APs, and removing also intersecting APs, thus resulting in 400 APs (221 real APs, 179 selected high traffic APs).

The resulting OD matrix for these distributions of APs has been estimated by computing the intersections between the real cab measures with the placed APs, according to a capturing range of 25 m radius. The OD matrix for this configuration was generated by evaluating the traffic counts among the various APs, grouped by the zip code they belong to. The chord diagrams of these scenarios are reported in Figure 6.

The interactive versions of the chords diagrams in which it is possible for each couple of locations to see the effective flows (in a way and in the other, for a given time slot of the day) are accessible on http://www.disit.org/6694.

IV. Experimental Results Analysis

A comparative analysis of traffic flows was conducted, using the above cited set of cab traces, consisting of 11,219,955 unique detections from 536 cabs, with respect to the above described scenarios. With the above assumptions the real set of APs placed in the city centre was used to sample the original data set, by calculating the APs intersections with the cab traces. The OD matrix was calculated from the sampled data set (considering each city zip code as a separate area), reporting the traffic counts among every city area.

This procedure was repeated by choosing the APs with a pseudo random technique, and by placing the APs only in the roads with the biggest amount of traffic. After that, a comparative statistical analysis was conducted for each configuration (see Table I). The traffic flow outcome is predicted with a linear regression, finding the parameters that best fit the data in the linear model

$$y = \alpha x + \beta \qquad (3)$$

where x is the dependent variable or predictor (i.e., traffic counts as registered by the sensors), and y is the outcome (i.e., predicted traffic counts). Building

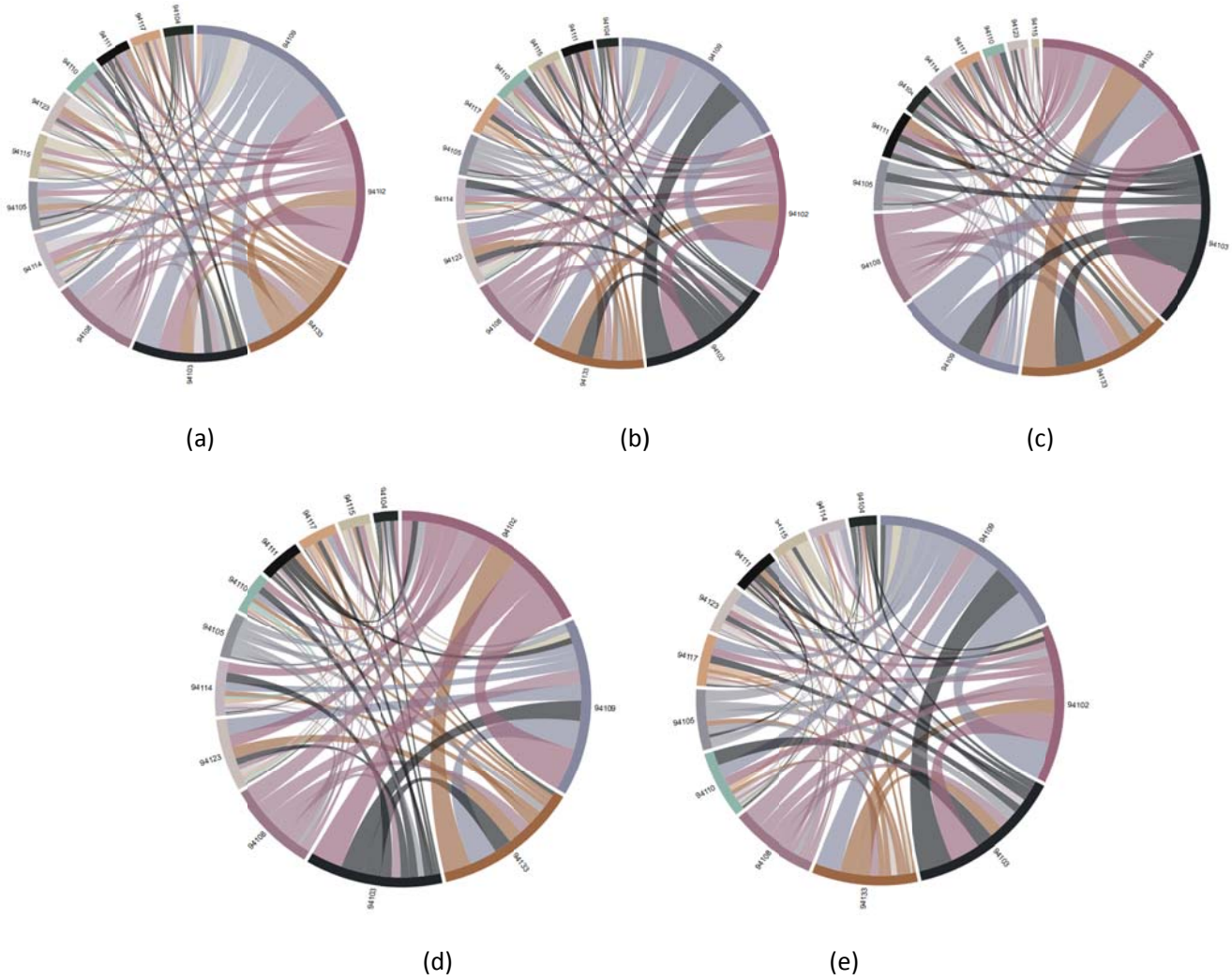


Figure 6. Chord diagram of flow counts. Cases as described in Table I: (a) Random APs; (b) High traffic APs; (c) High traffic APs (zip boundaries); (d) High traffic APs (top 400); (e) Real augmented APs

the model in (3) using the set of real APs gives a correlation of 0.446 (0.120 with real APs in the city downtown) with respect to the real traces. A number of cases have been assessed following the placement strategies described in Section III. In case (b), the APs have been placed on the roads with the highest traffic rate, producing a model with a correlation of 0.915, and of 0.835 using only the top 400 APs, as described in case (d); using random APs of case (a) gives a correlation of 0.835; using the APs only within 300 m from the areas' boundaries, described in case (c), gives a correlation of 0.687. It is clear from this data that using the real APs set produces noise and doesn't give a reliable model for the traffic counts. Randomly Distributing the APs gives a better correlation with the cab traces, while reducing the number of APs and considering only those in the proximity of each area, gives a good correlation while maintaining a limited number of APs. The set of real APs of case (e), integrated with some other APs and cleaned up from some unuseful or redundant elements (i.e., mutually intersecting APs), gives a correlation of 0.892). To visualize the results of the various OD models a web interface was developed, with the possibility to view the chord diagram for each computed configuration, for different time intervals of the day. This approach allowed to identify which are (i) the positions of the new APs to be added (i.e., 179) and (ii) the minimum set of APs already in place that have to be used for data acquisition (i.e., 229). The second point allows keeping limited the network bandwidth and the workload for the estimation of the OD matrix.

V. Conclusions

In this paper, we presented a flow analysis using OD models of traffic counts. Considering public Wi-Fi access points is not reliable in determining an effective model for the city traffic flows. Instead, placement of sensors only in proximity to the boundaries of high traffic routes yields a statistically significant linear relationship, between the OD matrix trip of sampled and real data. Thus, it is possible to model the traffic behaviour with a limited amount of traffic counters (in this context Wi-Fi access points) with an acceptable precision. The proposed methodology is general and can be applied to different urban scenarios, in the context of Smart City traffic management. It makes use of Wi-Fi access points (AP) distributed across the city of San Francisco. Comparative analysis has shown an increased precision of the AP positioning validated by obtaining an OD matrix and assessing correlation with the actual OD taken from the original real data by considering cab flows from high traffic zones, positioned in proximity of the zones boundaries, with respect to random uniformly distributed or real APs (e.g., restaurants, hotels, schools).

The approach allows identifying which are the strictly needed AP to be added with respect to the AP that can be already in place in the city, to exploit the whole infrastructure of Wi-Fi also for people flow monitoring and assessment.

ACKNOWLEDGMENT

The authors want to thank all the partners involved in the project Resolute (http://www.resolute-eu.org) and the European Commission for funding the project. Resolute has received funding in the scope of the Horizon 2020 Programme under Grant Agreement n. 653460 and in project Sii-mobility. The same approach is now used for instrumenting Florence, Italy.

REFERENCES

[Danalet et al., 2012] Danalet, Antonin, Michel Bierlaire, and Bilal Farooq. ”Estimating Pedestrian Destinations using Traces from WiFi Infrastructures.” Pedestrian and Evacuation Dynamics 2012. Springer International Publishing, 2014. 1341-1352.

[Patil et al., 2015] Patil, P. H., and A. A. Kokil. ”WiFiPi-Tracking at mass events.” Pervasive Computing (ICPC), 2015 International Conference on. IEEE, 2015.

[Gigginger et al., 2007] Giffinger, Rudolf; Christian Fertner; Hans Kramar; Robert Kalasek; Nataa Pichler-Milanovic; Evert Meijers, Smart cities Ranking of European medium-sized cities, 2007 Available at: http://www.smartcities.eu/

[Caragliu et al., 2009] Caragliu A., Del Bo C., Nijkamp P., Smart cities in Europe, 3rd Central European Conference in Regional Science CERS, 2009

[CARS 2020] Report on the state of play of the outcome of the work of the High Level Group, European Commission, October 2014

[Cascetta et al., 2001] Cascetta, E. and Postorino, M. N. (2001) Fixed point approaches to the estimation of O/D matrices using traffic counts on congested networks, Transportation Science 35: 134-147

[Bellini at al., 2014] Pierfrancesco Bellini, Monica Benigni, Riccardo Billero, Paolo Nesi, Nadia Rauch, Km4City ontology building vs data harvesting and cleaning for smart-city services, Journal of Visual Languages & Computing, Volume 25, Issue 6, December 2014, Pages 827-839

[Doblas et al., 2005] Doblas, J. and Benitez, F. G. (2005) An approach to estimating and updating origin-destination matrices based upon traffic counts preserving the prior structure of a survey matrix, Transportation Research, Part B: Methodological 39: 565-591

[Kim et al., 2001] Kim, H., Beak, S. and Lim, Y. (2001) Origin-destination matrices estimated with a genetic algorithm from link traffic counts, Transportation Research Record 1771, Transportation Research Board, Washington, D.C.:156-163

[Lundgren et al., 2008] Lundgren, J. T. and Peterson, A. (2008) A heuristic for the bilevel origin-destination matrix estimation problem, Transportation Research, Part B: Methodological 42: 339-354

[Nie et al., 2005] Nie, Y., Zhang, H. M. and Recker, W. W. (2005) Inferring origin-destination trip matrices with a decoupled GLS path flow estimator, Transportation Research, Part B: Methodological 39: 497-518

[Gong, 1998] Gong, Z. (1998) Estimating the urban o-d matrix: a neural network approach, European Journal of Operational Research 106: 108-115

[Ashok and Ben-Akiva, 2000] Ashok, K. and Ben-Akiva, M. E. (2000) Alternative approaches for real-time estimation and prediction of timede-pendent origin-destination flows, Transportation Science 34: 21-36

[Piorkowski et al., 2009] Piorkowski, M.; Sarafijanovic-Djukic, N.; Grossglauser, Matthias, ”A parsimonious model of mobile partitioned networks with clustering”, Communication Systems and Networks and Workshops, 2009. COMSNETS 2009. First International, pp.1,10, 5-10 Jan. 2009

Visualizing Transportation Routes for Data Analysis in Logistics

Paolo Buono
Dipartimento di Informatica
Università degli Studi di Bari Aldo Moro
Bari, Italy
paolo.buono@uniba.it

***Abstract*—Logistics activities refer to transporting materials and/or storing them in specific warehouse (platforms) for a given period of time. Typical problems are related to route and vehicle load optimization and monitoring of transportation conditions. GPS availability in the vehicles allows transportation companies to remotely check if vehicles are on schedule. Such information is often related to a single shipment. This paper presents a web tool which aims to support analysts through an interactive visualization technique that shows all routes performed by the vehicles in a given period of time on a geographic map. The tool allows analysts to perform explorative analysis in order to obtain information that otherwise is difficult to get. Various filters help them to reduce the visualized data to a manageable quantity. The interface is simple, complex queries can be executed by specifying a few parameters and performing zoom & pan gestures on the map.**

Keywords: route visualization; transportation analysis; logistics

I. Introduction

A recent project we worked to aimed at providing commercial opportunities to small and medium size enterprises by developing a system that enables all stakeholders involved in logistics to improve the cooperation and the visibility in the market. The project covered industrial processes from the acquisition of raw materials to the delivery of the final product to the client. There is a difference between the products and the materials to be transported. While the products are the result of a production process, transported materials also include intangible goods, such as information flows (orders, invoices, marked data, etc.). The combination of physical and information flows is an added value of the project. The developed system addresses a number of companies with different capability involved in logistics, which offer various goods and services. The system offers multimodal services and use new technologies to perform advanced planning, schedule, tracking, and analysis. This allows the producers, but also transportation companies, to reduce the quantity of unsold products, the storage time, to shorten or speed up the supply chain.

In early 50s the term logistics focused only on the movement of goods. Starting from the mid-1970s, methods to move goods more efficiently were considered. Today logistics includes the logical flow that accompanies the physical flow, which consists in documents and information of the process to pack, load, transport and deliver goods. Company managers that organize their fleets need to know the position of each vehicle, information about the route, transportation conditions and drivers' conditions. Vehicles are equipped with GPS. Sensors are often installed in the containers to constantly monitor the transport conditions. Today, transportation tracking is very common and the carrier can provide the client with the position or the stage in the process where the ordered product is.

On the carrier side, typical problems involve the optimization of vehicle load and routing, which often deal with mathematical models from the operational research field. About twenty years ago, routing was the main issue [1], then further issues related to load optimization and route scheduling occurred [2]. Later, with the goal of improving service quality and reducing costs, research focused on the planning of the entire supply chain [3]. Such methods are now robust and work pretty well. However, operational research algorithms are considered as black boxes, and the analysts do not get a precise understanding of how data are processed and the results are obtained. Moreover, optimization algorithms often take into account a single delivery and provide weak support for long-term analysis.

Common problems in logistics are due to: the high number of actors in the supply chain; the request of a small quantity of products; routing scheduling; load optimization; transportation condition quality; little or no use of intermodality. The developed system provides the involved companies with a time-based strategy [4] that performs appropriate planning and scheduling; allowing them to easily cooperate.

Logistics processes generate huge and dynamic datasets, that analysts find difficult to understand. The tool presented in this paper provides a visualization technique useful for analysts working in transportation, interested in understanding the complex and dynamic spatial interactions among products (either components or finished ones), producers, carriers (e.g. road, rail marine, flight companies, logistic providers), consignees, and clients (those who order goods and trigger transportations). The visualization technique has been developed according to the Shneiderman's Visual Information Seeking Mantra [6], which prescribes to provide a general overview of the data, but also the possibility to filter data in order to allow the user to concentrate on specific data of interest and, finally, to provide details on demand. Thus, the tool first shows an overview of routes of many shipments and provides filters to refine the visualized data in order to support the data analysis process. The user activates filters by modifying the value of

DOI reference number: 10.18293/DMS2016-040

various attributes, related to tracking (e.g. order date, delivery time, customer code), vehicles (e.g. transportation company, vehicle code, vehicle characteristics), load (e.g. weight, packaging, items no., product type), and delivery (e.g. address and location, date and time). Details of specific item are provided if the user requests them.

Next section briefly reports some related works about visualization of mobility, traffic, space and time, including also time series and process visualization. Section III presents the developed tool. Section IV concludes the paper, also highlighting some future work.

II. Related Work

To the best of our knowledge, logistics companies do not use advanced visualization to analyze data. They often use software systems whose outputs show icons overlaid on geographic maps, in order to indicate the position of a vehicle and provide information such as the status of vehicles (stopped, moving, broken, …).

Many works in the literature focus on mobility data, which refers to objects that change position during time. Movements can be represented as continuous paths in space that are also movement in time. Such continuous paths can also be seen as a composition of various spatial events [7]. A given path can also be enriched with additional data, such as the carrier and the cost for the shipment.

About traffic data visualizations, Chen et al. presented a survey that illustrates basic concepts and focuses on methods that address temporal, spatial, numerical and categorical properties [8]. In the survey, data are classified as location-based (where an object is detected), activity-based (what the object is doing), device-based (what is generating the data). The survey focuses on trajectories and incident visualization.

Time dependent data can be visualized in different ways, according to analysis goals and data characteristics. For instance, a visualization can be appropriate if the periodicity is relevant (in case anthropic activities or seasonal factors are important), while a completely different visualization is needed if time is considered as a continuous entity. Several proposals consider the periodicity, identified with radial [9] or spiral [10] visualizations. Many others visualize time linearly, often using line charts, which, however, are not very good in dealing with multivariate data. A survey on the most well-known visualization techniques for time-dependent data has been performed in [11]. The problem with multivariate visualization has been addressed in various ways. As an example, TimeSearcher adopts multiple panels, stacked and aligned according to a timeline in order to highlight correlations of the studied phenomena [12].

Abstract representations are well suited for visualizing process; the technique in [13] uses block diagrams to visualize supply chain data.

Many visualizations proposed in the literature have some issues that prevent utilization by potential users. Russell performed a study on a company that uses big data every day in order to improve its service [14]. The goal was to understand how such proposed visualizations are used today for analytical purposes. This study showed that new Information Visualization techniques are perceived useful in the investigation phase of research, but they appear to be peripheral to everyday work in companies. Among the reasons of such low adoption in working practice, there is the difficulty of preparing data for the visualization. This suggests that designers and developers should produce visualizations that, rather than being appealing, first and foremost are easy to understand and easy to use by the intended users. Even with simple representations, it is possible to perform good analyses. The example of the data table presented in [14] is very simple but it is well suited for the analysis it has to support.

III. The visualization tool

The goal of the work reported in this paper is to produce a simple, easy and immediate analytical tool that can display transportation information very quickly in order to support data analyst in their daily work. According to the Shneiderman's mantra [6], the tool provides an overview of available transportation routes on a geographic map, a set of filters to reduce the dataset to the item of interest and details of the visualized data, if required. As a case study, this paper presents data about shipments of a company in South Italy in 2014.

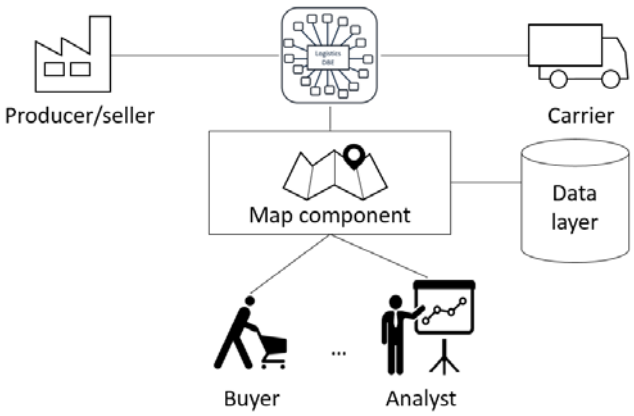


Figure 1. Simplified schema of the tool architectural organization.

Figure 1 shows a simplified schema of the tool architectural organization. Stakeholders interacting with the tool are represented at the bottom part of the schema, for example, a customer that buys a product and wants to keep track of the order, or an analyst who wants to find information about deliveries. *Map component,* a module that is part of *Logistic DBE*, allows the users to interact with the map and data represented on it. The map is visualized using OpenLayers libraries [5]. The *Logistic DBE (Digital Business Ecosystem)*, represented in the middle at the top of the schema of Figure 1, is the overall logistics system developed in the project, which stores information about logistics stakeholders, such as customers, carriers, producers, and service providers, for various activities. The DBE consists of many modules; its complete description is out of the scope of this paper. One of the core components is the EDI (Electronic Data Interchange) that allows various components to communicate in different formats, means, and modalities.

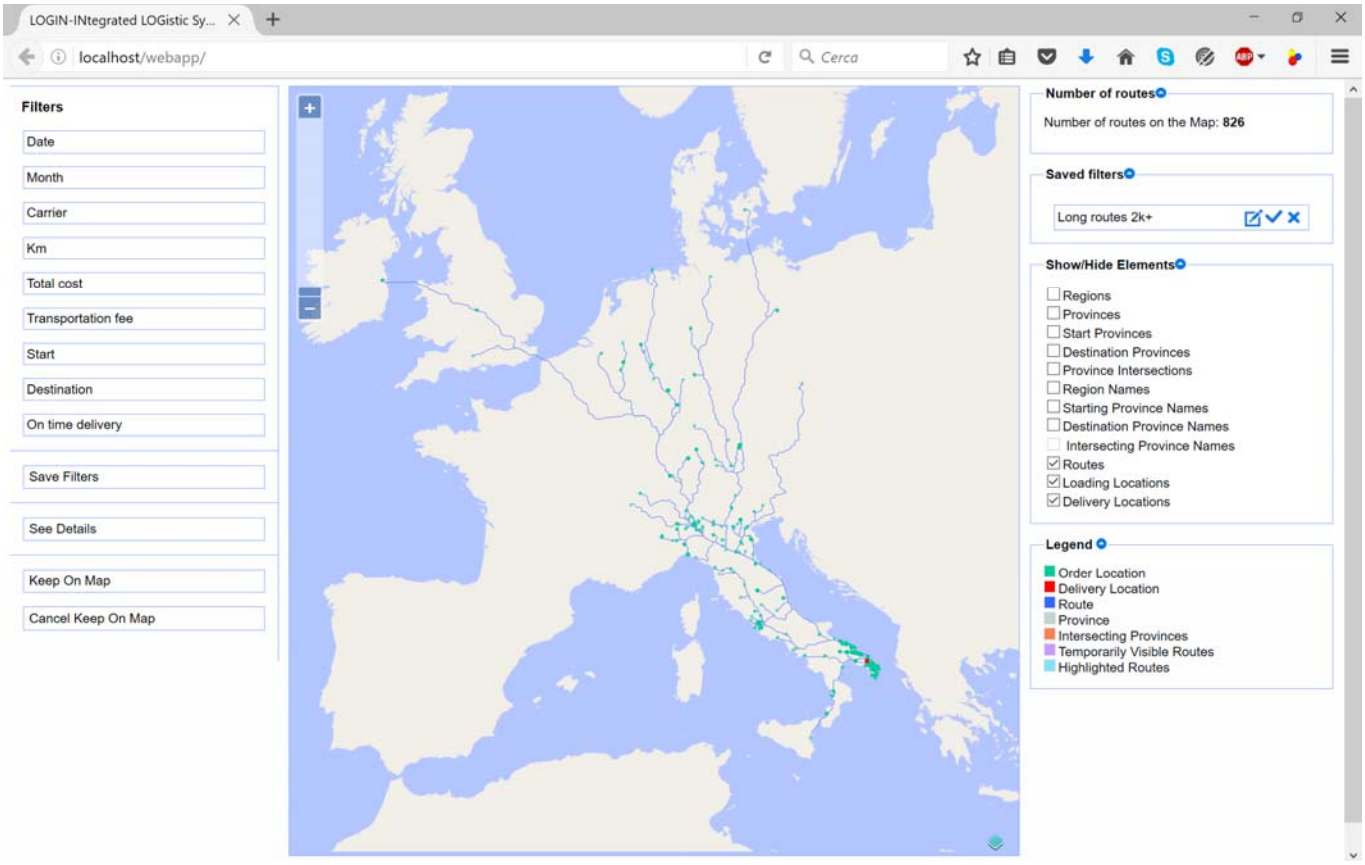


Figure 2. Overview of the transportation paths. A number of filters allow the analyst to select needed information.

Carrier (Figure 1) indicates companies that provide the service of shipping goods from the producer to the customer. Each vehicle sends information about the position and metadata of the transportation, in order to allow vehicle and good tracking. Vehicles send to the *Logistics DBE* data such as vehicle ID, timestamp, geographic coordinates, and data related to the sensors on board of the vehicle. Such sensor measurements can be related to temperature, solar radiation, humidity and other parameters relevant for the transported goods.

Producer/seller are the companies that offer products to the market and ask logistics solutions to the Digital Business Ecosystem.

One of the goals of Information Visualization is to help users forming their mental model about data. When dealing with geographic data and routing, it is quite obvious to show data through maps. Maps are used in transportation companies to have a clear idea about the position of their vehicles. During the project, we made on-site visit to transportation companies. We found that they use different tools for monitoring real-time conditions and traditional visual techniques are used to analyze historical data. An example is the use of histograms to visualize the time spent by a given driver on a vehicle. The analysis of the routes taken by the companies, however, were performed through specific queries to their database, with no support of any visual tool. The goal of the tool described in this paper is to help analysts to amplify their knowledge on data. Route data and layers for the representation of countries, borders, cities and other geographic information are stored in the *Data layer* (Figure 1).

The example reported in Figure 2 relates to the routes of different transportation companies that operated for a company that sold wine in the year 2014. Areas served by such company are immediately visible; in fact, the wine company works very much in South Italy, more specifically in the Apulia region. Many shipments go to Rome, Milan and other cities of Italy. The company has clients also in Europe and the routes mostly spread vertically leading to North. It is immediate to see that Spain and France are not part of the sale market of the company. The company does not even sell products to Sardinia and Corsica islands and performed just one single delivery to Sicily, namely to Catania.

On purpose, the map has a light solid colour, in order to avoid distracting the analyst with useless details. Region names and other details are shown on demand.

As visible in Figure 2, the user interface is divided in three main parts: the left panel contains a list of filters and operations to perform on selected data; in the middle, the map is shown; the right panel contains visualization options and the legend.

Available filters depend on the attributes in the company dataset (see left side of Figure 2). In particular, *Date* allows the user to select an interval of dates of interest. If the user wants to see all shipments that occurred, for instance, in March and in September, he or she should use the *Month* filter that allows

him/her to select noncontiguous months. *Carrier* is a searchable list of companies that provide the transportation service. Once selected, the user types a few character of the company name and the list narrows down, according to the typed characters. More than one transportation company can be selected. Figure 3 shows the selection of *Mail Boxes* and *BRT SpA*; consequently, all routes performed by other companies are removed from the visualization.

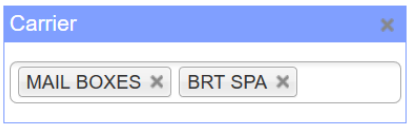


Figure 3. The Carrier filter shows the selection of two companies

Km filters all routes that satisfy the specified minimum and maximum distance between departure and destination. To set such values, the double slider shown in Figure 4. A double slider component filters all routes whose lengths are between 400 and 1320 Km. is used; in the shown example, the user selected all routes whose length is between 400 Km and 1320 Km. The extremes of the interval depend on the available dataset. In the case reported in Figure 4, the maximum distance is the destination to the Dublin area (Ireland), which is 2840 Km long. The use of a double slider is also appropriate for *Total cost* and *Transportation fees*, since they allow visualizing only routes whose total cost and transportation fees, respectively, are between minimum and maximum values specified by the user.

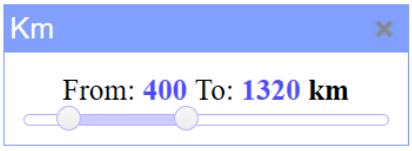


Figure 4. A double slider component filters all routes whose lengths are between 400 and 1320 Km.

Start and *Destination* filters are similar to the structure of *Carrier* filter, and are used to specify starting and destination locations. *On time delivery* filters only deliveries that are on time.

Save Filters saves the current filter combination for later use. Once selected, the user can leave the default value or he/she can assign a different name to easily remember it. In the example of Figure 2 there is one saved filter, available in the box *Saved Filters* on the right column and named by the user as "Long routes 2k+", meaning that the filter shows only routes longer than 2000 Km. When the filter is selected the visualization changes according to the parameters specified in the saved filter. The user may modify and delete the saved filters.

See Details is not a filter but once pressed, it shows the data in a tabular format. This is a common request made by the users, which do not completely believe in visualizations and want to see their data. It is also mentioned in the Shnederman's Mantra (details on demand) [6].

The user may also want to keep routes to make comparisons with other filter combination. *Keep on Map keeps* the current filter combination and shows the routes with another color in order to compare this with another selection. For instance, one may want to compare the routes made in the month of January with those of June. The button *Cancel Keep on Map* removes such temporary selection.

The *Number of Routes* box at the top of the right column in Figure 2 indicates the number of displayed routes.

As stated before, in order to avoid user overloading, the visualization initially shows the routes on the map with few details (see Figure 2). The user, depending on his/her needs, can operate on the parameters in order to see details of interest.

The *Show/hide Elements* panel on the right of the screen in Figure 2 contains buttons that give the possibility to visualize/hide elements on the map.

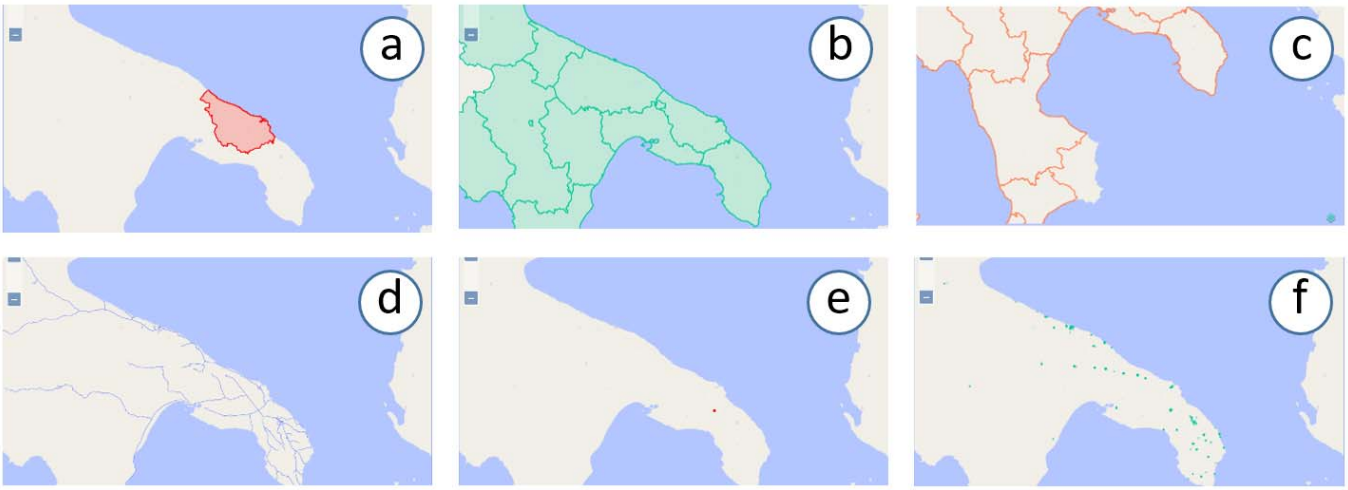


Figure 5. Different layers displayed on screen: a) province of the route starting point; b) provinces of destinations; c) provinces intersected by at least one route (routes are not visible); d) routes; e) route starting point (only one in this example) and destination cities.

Figure 5 shows some outputs generated after clicking on buttons in the *Show/Hide elements* panel: selecting the checkbox *Start Provinces* generates the output in Figure 5a, in which the starting locations (only one in this case) are shown; selecting the checkbox *Destination Provinces* generates the output in Figure 5b, in which the delivery locations are shown. selecting the checkbox *Province Intersections* generates the output in Figure 5c, in which the locations intersected by routes; selecting the checkbox *Routes* (selected by default) generates the output in Figure 5d, in which routes are shown; selecting the checkbox *Loading Locations* generates the output in Figure 5e, in which starting locations (only one in this case) of the routes are shown; selecting the checkbox *Delivery Locations* generates the output in Figure 5f, in which delivery locations of the routes are shown.

Figure 6 shows the visualizations of Figure 2 in which it is applied a zoom on Italy. With respect to Figure 2, several details are shown, including the borders of Provinces. The user can easily see that many provinces in Italy contain destinations with at least one delivery. Thirteen provinces that do not include destinations are crossed by routes. These areas should be further investigated in order to understand why there are no clients in such provinces. The dataset used for this example refers to a company that produces wine in the South Italy; it is evident that most of the clients are Italians. This company also sells products in Europe, even if there are very few delivery destinations (about 30, see Figure 2).

IV. Conclusions and Future Work

This paper presented a tool whose goal is to provide visualizations that help analysts of logistics data in understanding such data in order to provide a support for making decisions. The tool has been developed as part of a bigger system, composed by many modules, which aims to support companies in improving their product offer for their potential clients. In the design of the tool, Shneiderman's mantra [6] as well as suggestions to provide simple representations [14] were the guiding principles for generating a visualization that can be useful for its users.

The described tool could be enriched with other visualization techniques, which can reveal different aspects of data in order to increase user knowledge. One possibility can consider abstract routes instead of positioning them on the map. This could open to more possibilities for comparison, ordering and visualizing according to various attributes.

The tool has been developed on the basis of a User-Centred Design [15]. Informal tests with a few users were performed to evaluate the various tool prototypes. More formal user tests are planned in the future, involving domain experts; we expect to find indications in order to fine tune the present features and identify further useful feature to be implemented.

Acknowledgments

The author is grateful to G. Cucinelli for contributing to the development of the current version of the tool. This work is partially supported by the Italian Ministry of Economic Development (MISE) under grant PON Industria 2015 MI01_00294 "Login".

Figure 6. Various elements visualized on the Italy map

References

[1] L. Bodin and L. Levy, "Visualization in Vehicle Routing and Scheduling Problems," *ORSA Journal on Computing,* vol. 6, p. 8, 1994.

[2] A. Drexl and A. Kimms, "Lot sizing and scheduling — Survey and extensions," *European Journal of Operational Research,* vol. 99, pp. 221-235, 1997/06/01 1997.

[3] H. Meyr, J. Rohde, M. Wagner, and U. Wetterauer, "Architecture of Selected APS," in *Supply Chain Management and Advanced Planning: Concepts, Models, Software and Case Studies*, H. Stadtler and C. Kilger, Eds., ed Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 341-353.

[4] G. E. Spanner, J.P. Nuño, C. Chandra, Time-based strategies, Theory and practice, *Long Range Planning*, Volume 26, Issue 4, 1993, pp. 90-101, ISSN 0024-6301.

[5] OpenLayers. Available at: http://openlayers.org/. Last visit: Sept. 2016

[6] B. Shneiderman, "The eyes have it: a task by data type taxonomy for information visualizations", in Proceedings of the *IEEE Symposium on Visual Languages, 1996,* pp. 336-343.

[7] G. Andrienko, N. Andrienko, and M. Heurich, "An event-based conceptual model for context-aware movement analysis", *Int. J. Geogr. Inf. Sci.,* vol. 25, pp. 1347-1370, 2011.

[8] W. Chen, F. Guo, and F. Y. Wang, "A Survey of Traffic Data Visualization," *IEEE Transactions on Intelligent Transportation Systems,* vol. 16, 2015, pp. 2970-2984.

[9] J. Pu, S. Liu, Y. Ding, H. Qu, and L. Ni, "T-Watcher: A New Visual Analytic System for Effective Traffic Surveillance," in Proceedings of the *14th IEEE International Conference on Mobile Data Management*, 2013, pp. 127-136.

[10] J. V. Carlis and J. A. Konstan, "Interactive visualization of serial periodic data", in Proceedings of the *11th annual ACM symposium on User interface software and technology*, San Francisco, California, USA, 1998, pp. 29-38.

[11] W. Aigner, S. Miksch, H. Schumann, and C. Tominski, *Visualization of Time-Oriented Data*: Springer Publishing Company, 2011.

[12] P. Buono, C. Plaisant, A. Simeone, A. Aris, G. Shmueli, and W. Jank, "Similarity-Based Forecasting with Simultaneous Previews: A River Plot Interface for Time Series Forecasting", in Proceedings of the *11th International Conference Information Visualization, IV '07*, 2007, pp. 191-196.

[13] P. Buono, A. L. Simeone, C. Ardito, and R. Lanzilotti, "Visualizing data to support tracking in food supply chains", in Proceedings of the *15th International Conference on Distributed Multimedia Systems*, San Francisco Bay, CA, USA, 2009, pp. 369-374.

[14] D. M. Russell, "Simple is Good: Observations of Visualization Use Amongst the Big Data Digerati" in Proceedings of the *13th International Working Conference on Advanced Visual Interfaces, AVI 2016*, Bari, Italy, 2016, pp. 7-12.

[15] ISO/IEC 9241-210, 2010. Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems.

Authors' Index

Amato, Flora	1
Anjum, Bushra	69
Ardito, Carmelo	129
Bansal, Arvind	48
Bao, Peng	77
Belbachir, Hafida	98
Bellini, Pierfrancesco	202
Bouattou, Zina	98
Brylow, Dennis	112
Buono, Paolo	210
Case, Denise	105
Case, Denise M.	175
Casillo, Mario	165
Cenni, Daniele	202
Chang, Shikuo	119
Che, Xiaoping	61
Coccoli, Mauro	153, 158
Colace, Francesco	165
Costabile, Maria Francesca	129
Costagliola, Gennaro	16
Cremers, Armin B.	148
Cuzzocrea, Alfredo	84
De Rosa, Mattia	16
De Santo, Massimo	165
Desolda, Giuseppe	129
Dong, Honghui	55
Dong, Tiansi	148
Eloe, Nathan	9, 105
Eloe, Nathan W.	175
Favetta, Franck	91
Fish, Andrew	16
Fogli, Daniela	194
Francese, Rita	41
Fuccella, Vittorio	16
Ghayoumi, Mehdi	48
Grasso, Giorgio Mario	84
Gravino, Carmine	41

Hunpatin, Omokolade 112

Jia, Limin 184
Jiang, Yibin 148
Jin, Ke 77

Köhler, Joachim 148

Laurini, Robert 91, 98
Lemma, Saverio 165
Leopold, Jennifer 9, 105
Leopold, Jennifer L. 175
Liu, Weibin 138
Lombardi, Marco 165
Lu, Wei 61

Maresca, Paolo 34, 158
Matera, Maristella 129
Molinari, Andrea 34
Moscato, Vincenzo 1
Mumolo, Enzo 84

Nesi, Paolo 202

O'Hare, Casey 112

Peroni, Matteo 194
Pesare, Enrica 26
Picariello, Antonio 1

Qin, Yong 55, 184, 190

Risi, Michele 41
Roselli, Teresa 26
Rossano, Veronica 26

Sabharwal, Chaman 9, 69
Saleh, Rafiq 16
Scanniello, Giuseppe 41
Servigne, Sylvie 91
Sperlì, Giancarlo 1
Stanganelli, Lidia 158
Stefini, Claudia 194
Swartwood, Sarah 16

Thafar, Maha 48
Thomas, Ryan 112
Tortora, Genoveffa 41

Vercelli, Gianni 84, 153

Wang, Dandan 184
Wang, Li 190
Wang, Liqiang 61
Wang, Zhipeng 184

Xie, Zhengyu 190
Xing, Weiwei 61, 77, 138

Yang, Yanfang 55
Yang, Yong 61
Yu, Xiaomin 138

Zhang, Qing 55
Zhao, Xuejun 184

Program Committees' Index

Bilal Alsallakh	Vienna University of Technology
Ameer Armaly	University of Notre Dame
Michal Armoni	Weizmann Institute of Science
Timothy Arndt	Cleveland State University
Mario Arrigoni Neri	Università degli Studi di Bergamo
Dimitar Asenov	ETH Zurich
Danilo Avola	Sapienza University of Rome
Arvind Bansal	Kent State University
David Bau	MIT
Andrew Blake	University of Brighton
Paolo Bottoni	Sapienza University of Rome
Paolo Buono	Dipartimento di Informatica - Università degli Studi di Bari Aldo Moro
Loredana Caruccio	University of Salerno
Maiga Chang	Athabasca University
Shikuo Chang	Uinversity of Pittsburgh
Peter Chapman	University of Brighton
William Chu	Department of Computer Science and Information Engineering, TungHai University
Yuan-Sun Chu	National Chung Cheng University
Mauro Coccoli	DIBRIS - University of Genoa, Italy
Francesco Colace	University of Salerno
Luigi Colazzo	Università di Trento
Kendra Cooper	UT-Dallas
Gennaro Costagliola	Dipartimento di Informatica, Università di Salerno
Alfredo Cuzzocrea	ICAR-CNR and University of Calabria
Sergiu Dascalu	University of Nevada, Reno
Sayamindu Dasgupta	Massachusetts Institute of Technology
Andrea De Lucia	Department of Management & Information Technology - University of Salerno
Massimo De Santo	Università di Salerno
Aidan Delaney	University of Brighton
Vincenzo Deufemia	Department of Management & Information Technology, University of Salerno
Vincenzo Deufemia	Department of Computer Science, University of Salerno
Vincenzo Deufemia	University of Salerno
Georgia Dome	Wellesley College
Tiansi Dong	Bonn-Aachen International Center for Information Technology B-IT
Brian Dorn	University of Nebraska at Omaha
Michael Eisenberg	University of Colorado
Larbi Esmahi	Athabasca University
Filomena Ferrucci	Università di Salerno
Andrew Fish	University of Brighton
Scott Fleming	University of Memphis
Daniela Fogli	Università di Brescia

Rita Francese	University of Salerno
Mark Friedman	Google
Vittorio Fuccella	Dipartimento di Matematica e Informatica - Università di Salerno
Kaori Fujinami	Tokyo University of Agriculture and Technology
David Fuschi	Bridging Consulting Ltd
Ombretta Gaggi	Dept of Mathematics, University of Padua
Angelo Gargantini	University of Bergamo
Nikolaos Gkalelis	ITI
Mark Gross	University of Colorado
Angela Guercio	Kent State University at Stark
Philip Guo	University of Rochester
Brian Harvey	University of California, Berkeley
Bob Heller	Athabasca University
Carlos A. Iglesias	Universidad Politécnica de Madrid
Pedro Isaias	Universidade Aberta
Joaquim Jorge	IST/UTL/INESC-ID
Levent Kara	CMU
Andrew Ko	University of Washington
Jun Kong	North Dakota State University
Yau-Hwang Kuo	National Cheng Kung University
Robert Laurini	INSA Lyon
Jennifer Leopold	Missouri University of Science & Technology
Clayton Lewis	University of Colorado
Alan Liu	National Chung Cheng University
John Maloney	CDG Labs
Fred Martin	University of Massachusetts Lowell
Yoshiaki Matsuzawa	Aoyama Gakuin University
Xuelei Meng	Beijing Jiaotong University
Luana Micallef	Helsinki Institute for Information Technology HIIT
Nikolay Mirenkov	University of Aizu
Andrea Molinari	University of Trento
Jens Mönig	Y Combinator Research (HARC)
Max North	Southern Polytechnic State University
Yoshiki Ohshima	Communications Design Group, SAP Labs/Viewpoints Research Institute
Paolo Maresca Paolo Maresca	University Federico II
Ignazio Passero	Dipartimento di Matematica ed Informatica, Università degli Studi di Salerno, Italy
Marian Petre	Open University
Joseph Pfeiffer	New Mexico State University
Antonio Piccinno	University of Bari
Giuseppe Polese	University of Salerno
Yong Qin	Beijing Jiao Tong University
Elvinia Riccobene	DTI - University of Milan
Michele Risi	University of Salerno
Peter Rodgers	University of Kent
Teresa Roselli	Department of Computer Science - University of Bari

Eric Rosenbaum	Google
Veronica Rossano	Depatment of Computer Science - University of Bari
Krishnendu Roy	Valdosta State University
Natalie Rusk	MIT Media Lab
Giuseppe Scanniello	Dipartimento di Matematica, Informatica ed Economia
Monica Sebillo	Dipartimento di Management and Information Technology (DISTRA) - Università di Salerno
Ben Shapiro	University of Colorado, Boulder
Lidia Stanganelli	Università eCampus
Gem Stapleton	University of Brighton
Steve Tanimoto	University of Washington
Rob Thompson	
Franklyn Turbak	Wellesley College
Giuliana Vitiello	Dipartimento di Matematica e Informatica- Università di Salerno
Weiwei Xing	Beijing Jiaotong University
Zongyi Xing	Nanjing University of Technology and Engineering
Tangwen Yang	Beijing Jiaotong University
Atsuo Yoshitaka	Japan Advanced Institute of Science and Technology
Tomas Zeman	Czech Technical University in Prague

External Reviewers' Index

Cassano, Fabio
Di Nucci, Dario
Palomba, Fabio
Thompson, Robert
Tsai, Wen Hao

DMSVLSS2017 Call For Papers

The Twenty Third International Conference on Distributed Multimedia Systems, Visual Languages and Sentient Systems

Wyndham Pittsburgh University Center, Pittsburgh, USA
July 7 - 8, 2017
ksiresearchorg.ipage.com/seke/dmsvlss17.html

Organized by
KSI Research, USA

The main theme of the 23rd International Conference on Distributed Multimedia Systems, Visual Languages and Sentient Systems (DMSVLSS2017) is *Sentient Multimedia Systems*, which are distributed systems capable of actively interacting with the environment by gathering, processing, interpreting, storing and retrieving multimedia information originated from sensors, robots, actuators, websites and other information sources.

The conference organizers seek contributions of high quality papers, panels or tutorials, addressing any novel aspect of sentient multimedia systems that significantly benefits from the incorporation/integration of multimedia data (e.g., visual, audio, pen, voice, image, etc.), especially in visual languages and distance education technologies, for presentation at the conference and publication in the proceedings. Both research and case study papers or demonstrations describing results in research area as well as industrial development cases and experiences are solicited. The use of prototypes and demonstration video for presentations is encouraged.

Selected Papers of DMSVLSS2017 will be published in special issues of JVLC and VLSS

Co-Located Conference
The International Conference on Software Engineering and Knowledge Engineering (SEKE2017) will be held at the same hotel from July 5 to July 7, 2017. Authors are welcome to contribute papers and attend both conferences. Its website is: ksiresearchorg.ipage.com/seke/seke17.html

Topics include but are not limited to:

Distributed and Sentient Multimedia Technology
Distributed and Sentient Multimedia Models and Systems
Applications of Distributed and Sentient Multimedia Systems
Visual Languages and Applications
Distributed Education Technologies

DOI reference number: 10.18293/DMS2016-015

Workshops and Special Sessions
This year, DMSVLSS2017 will be held in conjunction with workshops and special sections. Papers submitted to workshops or special sessions are invited by the program committee chairs of the workshops/special sessions.

Hotel Information
The hotel has made available for these limited dates (7/4 - 7/9/2017) to DMSVLSS 2017 attendees a discount rate of $109 US dollars per room, per night, single/double occupancy, including free WiFi in hotel room and free parking, not including sales tax. For other days prior to or after the conference period, the hotel will provide best available rate. Guests can make on-line reservations by visiting the hotel's reservations website at www.wyndham.com. Please enter the hotel's zip code "15213" and arrival/departure date. Then click "special rates". Then click "group code". Enter the group code "07036841SE" and click "enter" and then finally click "find it". The name of the group and rate should appear. Please complete making the reservation in the usual manner. If you have problems please contact (412) 682-6200 and talk to Reservation Desk.

Information for authors
Papers must be written in English. An electronic version (PDF format) of the full paper should be submitted using the following URL:
https://www.easychair.org/conferences/?conf=dmsvlss2017. When submitting, select the main conference DMSVLSS2017 or one of the special tracks. Manuscript must include a 200-word abstract and no more than 8 pages of double column text (including figures and references). All submissions must not be published or under consideration for publication in a journal or in a conference with proceedings. Papers will be evaluated based on originality, significance, technical soundness and clarify of exposition. Depending upon the results of evaluation a paper may be accepted as regular paper (10 pages), short paper (7 pages) or technical notes (1 page and demo/poster presentation).

Information for reviewers
Papers submitted to DMSVLSS2017 will be reviewed online. The users (webmaster, program chair and reviewers) can login using the following URL:
https://www.easychair.org/conferences/?conf=dmsvlss2017.

Contact Information
DMSVLSS2017 Conference Secretariat
KSI Research
156 Park Square, Pittsburgh, PA 15238, USA
E-mail: dmsvlss2017@easychair.org

Important dates
Paper submission due: March 15, 2017
Notification of acceptance: April 20, 2017
Camera-ready copy of revised paper: May 10, 2017
Early conference registration due: May 10, 2017

Proceedings of the Twenty-Second International Conference on Distributed Multimedia Systems

and

Journal of Visual Languages and Sentient Systems, Volume 2, 2016

DMS

Salerno, Italy
November 25-26 2016

Copyright © 2016

Printed by
KSI Research Inc. &
Knowledge Systems Institute
Graduate School

156 Park Square
Pittsburgh, PA 15238 USA
Tel: +1-412-606-5022
Fax: +1-847-679-3166
Email: seke@ksiresearch.org
Web: http://ksiresearchorg.ipage.com/seke/dms16.html

ISBN 1-891706-40-3 (paper)
ISSN 2326-3261 (print)