Proceedings of the 2017 International Conference on Decision Support System Technology

Linden, Isabelle; Mareschal, Bertrand ; Liu, Shaofeng; Papathanasiou, Jason; Colot, Christian

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

Proceedings

of the

2017 International Conference on Decision Support System Technology

with a Theme on

Data, Information and Knowledge Visualisation in Decision Making

Namur, Belgium

29 - 31 May 2017

Editors:

Isabelle Linden, Bertrand Mareschal, Shaofeng Liu, Jason Papathanasiou, Christian Colot

ISBN :978-2-917490-28-0	**Copyright** © 2017 ICDSST. All rights reserved.

ICDSST 2017 Proceedings
Table of Content

ICDSST 2017 Short Papers 51

ICDSST 2016 Posters 185

Dean's Welcome Message

Faculty of Economics, Social Sciences and Business Administration
University of Namur

Professor Paul Wynants

On behalf of the Faculty of Economics, Social Sciences and Business Administration at the University of Namur, I'm delighted to welcome you all to Namur for the 2017 International Conference on Decision Support System Technology – ICDSST 2017.

It is now a commonplace to say that information sources are multiple, varied and diversified. However, the ability to navigate through this data, interpret it and extract sense require advanced capacities and continuously to develop and enhance them.

For years, the Business Administration Department of our faculty proposes advanced business engineering programs closely associating ICT and business skills. Decision Support Systems conception and development require highly skilled people having developed such a twofold profile.

It is a great honour to be hosting your prestigious conference this year. May this event offer you many enriching opportunities and fruitful debates.

I wish you a wonderful time at the conference in Namur!

Preface

This Proceedings presents the short papers, posters and abstracts of full papers accepted to the 3rd International Conference on Decision Support System Technology, ICDSST 2017, held in Namur, Belgium, during May 29th to 31st, 2017. This event had a main theme *Data, Information and Knowledge Visualisation in Decision Making.* This event is organized by the Euro Working Group on Decision Support Systems (EWG-DSS) in collaboration with the University of Namur Belgium, the EFFaTA Research Team of the University of Namur and the Université libre de Bruxelles, Belgium.

The main purpose of the ICDSST 2017 was to attract researchers, developers and specialists in the related areas of decision making, including its methodologies and technologies, as well as application oriented practitioners directed to the implementation of solutions for societal challenges, to get together in order to exchange experiences and identify the key issues within the areas in focus, so that exploitation of new approaches and tools for future developments can take place.

Depending on the research domain, frontiers between data, information and knowledge are not always very formally specified. Whatever the chosen definition, one will agree on the fact that the three of them are concerned with decision, decision making and decision support

Specific topics of interest of ICDSST 2017 include:

- *Qualities of Data Visualisation*
- *Data visualisation and Decision Making/Support*
- *Relational Data Visualisation and Decision Making/Support*
- *Innovative Data Visualisation, Manipulation and Interaction*
- *Social and Network Data Visualisation*
- *Textual Data Visualisation*
- *Qualitative Data Visualisation*
- *Process Visualisation*
- *Spatio-temporal Data Visualisation and Management*
- *Environmental Data Visualisation*
- *Visual Information Extraction and Navigation*
- *Geographic Information Systems and Decision Making/Support*
- *Health Information Visualisation*
- *Visualisation Supporting Knowledge Extraction*
- *Big Data Analytics*
- *Business Intelligence, Managerial Dashboard*
- *Knowledge Acquisition and Management*
- *Knowledge Extraction and Visualisation*
- *Knowledge Communication through Visualisation*
- *Visualisation and Collaborative Decision Making*
- *Decision Making in Modern Education*

This wide and rich variety of topic areas allowed us, on the one hand, to present a collection of innovative solutions to real decision making process in a range of domains, on the other hand, to highlight the main trends and research evolution in DSS. This Proceedings edited by EWG-DSS and Local Organisation Committee has considered contributions from a large number of submissions which were rigorously reviewed, maintaining EWG-DSS' long established reputation and standards of high quality. Each paper/ poster was reviewed by at least two internationally known experts from the ICDSST 2017 Program Committee comprising 58 scholars and practitioners from 22 countries. The Conference received in total 58 submissions. Following a two-stage review process and based on the review reports and recommendations from the Programme Committee, 13 full papers, 19 short papers and 9 posters were accepted. This ICDSST 2017 Proceedings includes the short papers, the posters and the abstracts of the full papers.

We would like to thank many people who have greatly helped the success of this Proceedings. First of all, we would like to thank the EWG-DSS to give us the opportunity to edit the Proceedings for ICDSST 2017. Secondly, we highly appreciate the help from University of Toulouse Library for providing us with the ISBN Number that is used for the publication of this Proceedings with a unique ID. Thirdly, we would like to thank the Faculty of Economics, Social Sciences and Business Administration at University of Namur for providing us a fantastic avenue and all necessary facilities for the Conference. Fourthly, we need to thank all the authors for submitting their state-of-the-art work to be considered for the Proceedings. All accepted papers and posters are of extremely high quality. Finally, we wish to express our gratitude to the reviewers, who have voluntarily helped with the evaluation and improvement of the papers and posters at their busiest time of the year.

We believe that this Proceedings has made a high-quality selection of well-balanced and interesting research papers and posters addressing the conference main theme. We hope that the readers will enjoy the publication!

The Editors:
Professor Isabelle Linden, University of Namur, Belgium
Professor Bertrand Mareschal, Université libre de Bruxelles, Belgium
Professor Shaofeng Liu, University of Plymouth, United Kingdom
Dr. Jason Papathanasiou, University of Macedonia, Thessaloniki, Greece
Mr. Christian Colot, University of Namur, Belgium

EWG-DSS

The EWG-DSS is a Euro Working Group on Decision Support Systems within EURO, the Association of the European Operational Research Societies. The main purpose of the EWG-DSS is to establish a platform for encouraging state-of-the-art high-quality research and collaboration work within the DSS community. Other aims of the EWG- DSS are to:

- Encourage the exchange of information among practitioners, end-users, and researchers in the area of decision systems

- Enforce the networking among the DSS communities available and facilitate activities that are essential for the start up of international cooperation research and projects

- Facilitate the creation of professional, academic, and industrial opportunities for its members

- Favor the development of innovative models, methods, and tools in the field of decision support and related areas

- Actively promote the interest on decision systems in the scientific community by organizing dedicated workshops, seminars, mini-conferences, and conference, as well as editing special and contributed issues in relevant scientific journals

The EWG-DSS was founded with 24 members, during the EURO Summer Institute on DSS that took place at Madeira, Portugal, in May 1989, organized by two well- known academics of the OR community: Jean-Pierre Brans and José Paixão. The EWG-DSS group has substantially grown along the years. Currently, we have over 300 registered members from around the world.

Through the years, much collaboration among the group members has generated valuable contributions to the DSS field, which resulted in many journal publications. Since its creation, the EWG-DSS has held annual meetings in various European countries, and has taken active part in the EURO Conferences on decision-making-related subjects. Starting from 2015, the EWG-DSS established its own annual conferences, namely, the International Conference on Decision Support System Technology (ICDSST).

EWG-DSS Coordination Board

Prof. Dr. Pascale Zaraté
Toulouse University /IRIT, France
E-Mail: zarate@irit.fr

Dr. Fatima C.C. Dargam
SimTech Simulation Technology & ILTC, Graz, Austria
E-mail: f.dargam@simtechnology.com

Prof. Dr. Rita Ribeiro
UNINOVA – CA3, Portugal,
E-Mail: rar@uninova.pt

Prof. Dr. Shaofeng Liu
University of Plymouth, UK
E-Mail: shaofeng.liu@plymouth.ac.uk

Prof. Dr. Isabelle Linden
University of Namur, Belgium
E-Mail: isabelle.linden@unamur.be

Prof. Dr. Boris Delibasic
University of Belgrade, Serbia
E-Mail: delibasicboris@gmail.com

Dr. Jason Papathanasiou
University of Macedonia,Thessaloniki, Greece,
E-Mail: jason.papathanasiou@gmail.com

Organizing Committee

Local Committee

Professor Isabelle Linden
(Local co-Chair)
University of Namur, Belgium,
Department of Business Administration

Professor Bertrand Mareschal
(Local co-Chair)
ULB, Université Libre de Bruxelles
Solvay Brussels School of Economics & Management

Mr Christian Colot
University of Namur, Belgium

Mme Pierrette Noël
University of Namur, Belgium

Mme Aurore Midavaine
University of Namur, Belgium

EFFaTA Team, University of Namur
Evocative Framework For Text Analysis – Mediality Models
https://effatamem.wordpress.com/

Prof. Isabelle Linden, Department of Business Administration,
Prof. Anne Wallemacq, Department of Business Administration,
Prof. Bruno Dumas, Faculty of Computer Science,
Prof. Guy Deville, Department of Modern Languages
Prof. Anke Bosse, German Department

EFFaTA-MeM

Evocative Framework For Text Analysis – Mediality Models

Anke Bosse, Anne Wallemacq, Bruno Dumas, Guy Deville, Isabelle Linden, Antoine Clarinval

Context

Text Analysis from the point of view of

Human sciences
- Qualitative
- Word vision
- Perception

Computer science
- Quantitative
- Knowledge extraction
- Cognition

Towards Cross-fertilisation

A Paradigm Shift

From text >< visualisation

To

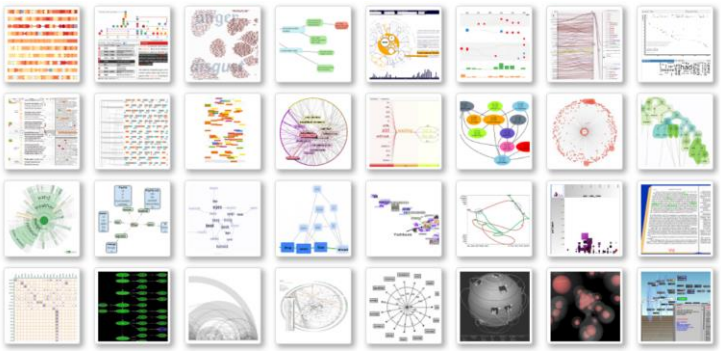
Text == visualisation

& Study of Intermedial Transposition

Text Spaces

Linear	• Architectural & Simultaneous
Univocal	• Plurivocal
2D	• 3D
Author >< reader	• Interactive & Cocreation
Frontal	• Immersive

Mainstream Visualisations of Text



Kostiantyn Kucher and Andreas Kerren, *Text visualization techniques: Taxonomy, visual survey, and community insights*, in Visualization Symposium (PacificVis), 2015 IEEE Pacific

Implicit Conception of Language

Univocal Semantics

Text reduced to a bag of words

or a set of relations between terms

In Search of a New Text Space

Semantic landscapes

Sensitive

Connotations

Immersive

Associations/oppositions

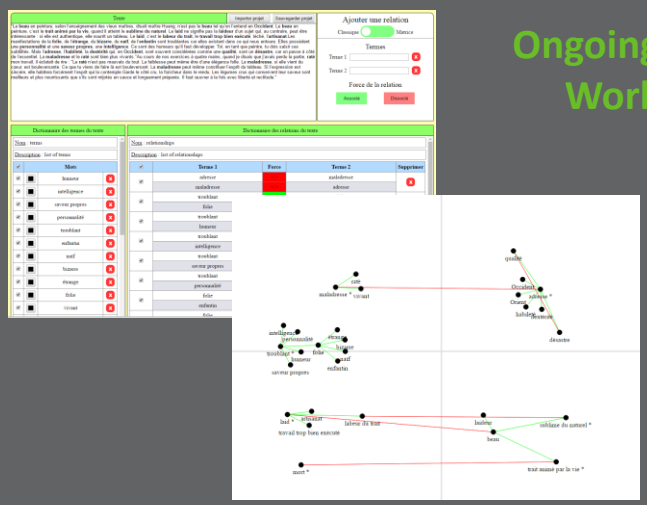
Word spaces

Polyvocal

Tactile

Towards a synesthethic experience

Ongoing Work



UNIVERSITÉ DE NAMUR

Programme Committee

Abdelkader Adla, University of Oran 1, Algeria

Adiel Teixeira de Almeida, Federal University of Pernambuco, Brazil

Alex Duffy, University of Strathclyde, UK

Alexander Smirnov, Russian Academy of Sciences, Russia

Alexis Tsoukias, University Paris Dauphine, France

Alok Choudhary, Loughborough University, UK

Ana Paula Cabral, Federal University of Pernambuco, Brazil

Ana Respício, University of Lisbon, Portugal

Andy Wong, University of Strathclyde, UK

Ben C.K. Ngan, Pennsylvania State University, USA

Bertrand Mareschal, Université Libre de Bruxelles, Belgium

Boris Delibašić, University of Belgrade, Serbia

Carlos Henggeler Antunes, University of Coimbra, Portugal

Csaba Csaki, University College Cork, Ireland

Dobrila Petrovic, Coventry University, UK

Dragana Bečejski-Vujaklija, Serbian Society for Informatics, Serbia

Fátima Dargam, SimTech Simulation Technology/ILTC, Austria

Festus Oderanti, University of Plymouth, UK

Francesca Toni, Ph.D. Professor at Imperial College London

Francisco Antunes, Beira Interior University, Portugal

François Pinet, Cemagref/Irstea, France

Frantisek Sudzina, Aalborg University, Denmark

Gabriela Florescu, National Institute for Research and Development in Informatics, Romania

Gloria Philipps-Wren, Loyola University Maryland, USA

Gregory Kersten, Concordia University, Canada

Guy Camilleri, Toulouse III University/IRIT, France

Hing Kai Chan, University of Nottingham, Ningbo Campus, UK/China

Irène Abi-Zeid, FSA – Laval University, Canada

Isabelle Linden, University of Namur, Belgium

Jan Mares, University of Chemical Technology, Czech Republic

Jason Papathanasiou, University of Macedonia, Greece

Jean-Marie Jacquet, University of Namur, Belgium

João Lourenço, Universidade de Lisboa, Portugal

João Paulo Costa, University of Coimbra, Portugal

Jorge Freire de Souza, Engineering University of Porto, Portugal

Jorge Hernández, University of Liverpool, UK

José Maria Moreno Jimenez, Zaragoza University, Spain

Kathrin Kirchner, Berlin School of Economics and Law, Germany

Lai Xu, Bournemouth University, UK

Leonilde Varela, University of Minho, Portugal

Lynne Butel, University of Plymouth, UK

Marc Kilgour, Wilfrid Laurier University, Canada

Marko Bohanec, Jozef Stefan Institute, Slovenia

Nikolaos Matsatsinis, Technical University of Crete, Greece

Nikolaos Ploskas, Carnegie Mellon University, USA

Pascale Zaraté, IRIT/Toulouse University, France

Pavlos Delias, Kavala Institute of Technology, Greece

Petraq Papajorgji, Canadian Institute of Technology, Albania

Priscila Lima, Instituto Tercio Pacitti – UFR Federal University of Rio de Janeiro, Brazil

Rita Ribeiro, UNINOVA – CA3, Portugal

Rudolf Vetschera, University of Vienna, Austria

Sean Eom, Southeast Missouri State University, USA

Shaofeng Liu, University of Plymouth, UK

Stelios Tsafarakis, Technical University of Crete, Greece

Uchitha Jayawickrama, Staffordshire University, UK

Institutional Support and Sponsors

Main sponsors

Working Group on Decision Support Systems
(http://ewgdss.wordpress.com/)

Association of European Operational Research Societies
(www.euro-online.org)

Institutional Sponsors

University of Namur, Belgium
(http://www.unamur.be/)

Fonds de la Recherche Scientifique, Belgium
(http://www1.frs-fnrs.be/index.php)

Graduate School of Management,
Faculty of Business, University of Plymouth, UK
(http://www.plymouth.ac.uk/)

University of Toulouse, France
(http://www.univ-tlse1.fr/)

IRIT Institut de Research en Informatique de Toulouse, France
(http://www.irit.fr/)

SimTech Simulation Technology, Austria
(http://www.SimTechnology.com)

Faculty of Organisational Sciences, University of Belgrade, Serbia
(http://www.fon.bg.ac.rs/eng/)

UNINOVA – CA3 – Computational Intelligence Research Group Portugal
(www.uninova.pt/ca3/)

University of Macedonia, Department of Marketing and Operations Management Thessaloniki, Greece
(http://www.uom.gr/index.php?newlang=eng)

Industrial Sponsors

Springer
(www.springer.com)

Lumina Decision Systems
(www.lumina.com)

ExtendSim Power Tools for Simulation
(http://www.extendsim.com)

Paramount Decisions
(https://paramountdecisions.com/)

1000 Minds
(https://www.1000minds.com/)

Reach Consultancy
(www.reach-consultancy.eu)

Professional Society Sponsors

IUFRO
(http://www.iufro.org/)

ICDSST 2017 Invited Speakers

Invited Speakers

Doctor Erica Yang
Technology Division, Scientific Computing Department,
STFC Rutherford Appleton Laboratory, Oxfordshire, UK
Talk: *Visualisation that underpins a new wave of visual computing technologies for big data analytics: methods and applications*

Professor Jean-Pierre Brans
ULB
Talk: *Ethics and Decision*

Professor Pierre Kunsch
VUB
Talk: *Ethics and Decision*

Doctor Abel Coll
CIMNE Barcelona
Talk: *Using Big Data for extreme large scale simulations analysis and visualization*

Dr Erica Yang

Dr Erica Yang is Head of Visualisation in the Scientific Computing Department at the Science and Technology Facilities Council (STFC) in the UK, and also an invited independent expert appointed by the European Commission for the H2020 programme. She has background in distributed system engineering, data science, and visualisation. Following a PhD in computer science from University of Durham, a post-doctoral research fellowship at University of Leeds, she joined STFC Rutherford Appleton Laboratory in 2007 as a senior scientist. She was prompted to the head of visualisation position in 2016 to bring together two major strands of work undertaken by the Department: high throughput complex data analysis and visualisation. She has delivered high profile computational projects with UK's national science facilities and worked extensively with large laboratories in Europe and the US, in addressing data and compute intensive problems in big science experiments. She also has extensive collaboration links with civil engineering, automotive, and transportation industries, particularly in developing scalable visual analytics systems to tackle increasingly prevalent complex data analysis, visualisation, and exploration challenges in these domains. This talk will give an overview of the group's current projects with particular highlights of the role visualisation plays in a variety of decision making processes.

Talk : Visualisation that underpins a new wave of visual computing technologies for big data analytics: methods and applications

The rapid rise of big data is transforming many aspects of our society, from science and engineering in world class laboratories to many sectors which do not conventionally have substantial presence of computing technologies, e.g. construction engineering and agriculture. The volume, velocity, variety, and veracity (the 4Vs) and perhaps, more importantly, the complexity of data and the diversity of domains that data come from have meant that humans increasingly depend on the effectiveness of visualisation methods and systems. The main thrust of this talk is to argue that visualisation, whilst it can be a subjective matter, has become a critical tool for decision makers to make sense of big data & analytics methods. Applications range from high throughput image analysis that takes place in large laboratories to streaming analytics that look at the motions of social media and increasingly Cyber Physical Systems (CPS) and Internet of Things (IoTs). This talk will examine closely the uncertainty of visualisation methods through comparing and contrasting the differences among several popular visualisation frameworks and tools, e.g. D3 for semantic text analytics, to ParaView for large scale volume rendering and in-situ visualisation for computational flow simulation. Finally, this talk covers the increasingly popular array of high end computing methods and parallel computational infrastructure that underpin the new wave of visual analytics capabilities which are essential for big data systems to deliver high throughput analytical insights from computational modelling, simulation and prediction.

Jean-Pierre Brans

Jean-Pierre Brans received his Ph.D. in Mathematics at the ULB/VUB University of Brussels (1966). He has been Professor in these Universities since 1964 and dean of the VUB-Solvay Business school (1975-78). He has held visiting professorships at the universities of Kinshasa, Constantine, Aix-en-Provence, ENST (Brest), ENSTA (Paris), AIT (Bangkok), KUL (Leuven), and Lulea (Sweden). He has taught on statistical analysis, OR, MCDA, mathematical programming and system dynamics. Prof. Brans has been president of the Belgian OR society (1975-78), president of EURO K Conference and organisor of the first one in Brussels (1975), initiator of the EURO Summer Institutes and organisor of the two first ones, initiator of the MINI EURO Conferences and organisor of the 1st, the 7th and the 12th ones. He has presented over 100 invited lectures all over the world and published over 100 papers in international scientific journals. He is the initiator of the PROMETHEE-GAIA Methodology and has written one book on this subject with B. Mareschal. He received teh EURO Gold Medal 1994, Doctor Honoris Causa from Copenhagen (2000) and Doctor Honoris Causa from Fribourg (2002). He has been Elected Professor of the year by the Students of the VUB University Brussels in 2002.

Pierre Louis Kunsch

Pierre Louis Kunsch born in 1947 has a Master in Physics from Université de Liège (Belgium, 1969), a PhD in Science from ETHZ Federal School of Technology Zürich (Switzerland, 1977), and a Higher Education Tenure in Operations Research (OR) from ULB Université Libre de Bruxelles in Brussels (Belgium, 1996).
His career combines technical, economic and OR studies in both industrial and academic environments.
Today he teaches OR and system dynamics at Université Libre de Bruxelles (ULB) and Vrije Universiteit Brussel (VUB) in Brussels. His main research domains are energy economics, multiple-criteria analysis, environmental economics, modelling and simulation of complex socio-economic systems with system dynamics and agent-based modelling.
He has published about 100 scientific articles and book chapters on these topics. He is past President and Delegate Administrator of ORBEL, the Belgian OR Society

Webbsites:

www.orbel.be
http://homepages.ulb.ac.be/~pikunsch/
https://cris.cumulus.vub.ac.be/portal/en/persons/pierre-kunsch(b0d7d137-a283-4229-8b62-ba8d91999b95)/publications.html
https://www.researchgate.net/profile/Pierre_Kunsch

Talk : Ethics and Decision

Although the pace of progress continues fast and in a positive way in today's world, many problems may arise because ethical issues are often completely ignored by decision makers. Ethics is a philosophical concept implying respect in all human, economic, social and environmental dimensions in a decision process. Its nature is dynamic, and this makes it a concern in an increasingly complex world.

In a first part of the talk an overview will be given on the evolution in the last decades in Operations Research in general and Decision Support Systems (DSS) in particular, moving from mono-criterion optimisation approaches to multiple-criterion decision-making. Though Ethics seems to be at the origin a rather abstract philosophical concept, the authors' tenet is that ethical values can and must be integrated into quantitative modelling. A methodology called **ASCM**, standing for **A**daptive, **S**ystemic, **C**ontrol and **M**ulti-criteria based methodology, is proposed to that aim by the authors to approach real-world problems. It includes multi-criteria analysis – here the well-known PROMETHEE methodology can be used in practice – incorporated into a wider dynamic decision-making process using systemic techniques – here in practice system dynamics or agent-based modelling. ASCM looks extremely promising for innovative DSS developments. It allows controlling and monitoring the evolution of human systems in a structured way, not ignoring important ethical aspects.

In a second part of the talk, examples of complex decision making processes in today's world involving important ethical dimensions will be presented within the ASCM framework.

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

Multi-criteria weight elicitation by multiple stakeholders piloting complex systems with ASCM

Pierre L. Kunsch, Jean-Pierre Brans
Vrije Universiteit Brussel, BUTO Dept.
Pleinlaan 2, BE-1050 BRUSSELS, Belgium
pkunsch@vub.ac.be, brans.jp@skynet.be
web-page: http://research.vub.ac.be/buto/members

ABSTRACT

In this short article an efficient weight elicitation technique is proposed for inclusion into the Adaptive, Systemic, Control and Multi-criteria based methodology, in short ASCM, presented during a plenary session of the ICDSST 2017 conference. The purpose of ASCM is piloting in real time complex systems by combining System Dynamics (SD) and Multi-Criteria Decision Aid (MCDA). Piloting policies are established and revised on a regular basis and/or constant real-time observation by means of SD simulations; groups of stakeholders (SH) choose by means of MCDA tools the best policy to be implemented for the ensuing time periods before possible adaptations to account for the actual system evolution. The weight elicitation process must not require excessive cognitive effort for the SH groups which might have different decisional powers; it must remain transparent, flexible, and allow easy sensitivity analysis for changes when confronting different opinions. Ideas from the well-known Simos-card procedure and ordinal rank-based weights were combined to a two step procedure defining firstly importance classes on ordinal Likert scales, and secondly profiles on those scales for the criteria. This two step approach gives a very flexible technique, at the same time transparent, easy to understand and to use for combining many different diverging opinions of an arbitrary number of SH groups. The article illustrates with an example those properties obtained for an efficient use of ASCM by multiple SH groups.

Keywords: Complex systems, ASCM methodology, System dynamics, Multiple criteria, Weight elicitation, Multiple stakeholders

INTRODUCTION

Weight elicitation is an essential piece of information for Multi-criteria Decision Aid (MCDA) [1], in particular for ASCM, standing for Adaptive, Systemic, Control and Multi-criteria based methodology, presented in this ICDSST 2017 conference by two of its authors [2, 3, 4]. The purpose of the present short paper is to introduce a weight elicitation methodology supplementing ASCM, considering that it must remain simple and transparent to all involved stakeholders.

ASCM is an interactive and iterative hybrid procedure combining MCDA tools with System Dynamics and Control Theory to pilot in real-time complex systems along sustainable paths. This tool has been developed specifically for the purpose of solving complex social-economic problems, and ethical dilemmas of which there are many in the contemporaneous world. Along a time path long-term policies are re-examined on a regular basis, or when needed by evolutions external to the system; simulation tools are used for that purpose, proposedly System Dynamics developing a strong feedback-view of the system under focus; the choice of a suitable multi-criteria policy is made by means of PROMETHEE [5], or another MCDA tool. Ethical decision-making implies next to multiple generally conflicting criteria the intervention of many different groups of stakeholders (SH) in the Golden Triangle of Economy, Social Life, and Environment [6], and this makes the choice of transparent criterion weights a central issue.

MCDA weights express the relative importance of criteria, but not all approaches include weight elicitation assistance; or if they do they often require much cognitive efforts, especially when different opinions and judgements of SH co-exist about the relative importance. The presence of ethical dilemmas, for example a subsidiary delocalising process today common in the global world, implicates two vertices of the Golden Triangle (Economy and Social Life), so that trade unions and financiers have contradictory points of view. An elicitation procedure must therefore have several transparencies and visibility features, as required by SH, as follows:

1. A full knowledge is required of the procedure applied to combine an arbitrary number of criterion importance judgements; the procedure must be accurate but remain simple; no black box is here acceptable, working mechanisms must be easy to grasp, also for laymen outside the field of decision science;
2. A clear view must be available on how possible differences of treatment between SH are accounted for, for example by how much a board member is given more decisional power than an executive director;
3. Full transparency regarding final elicitation results must be provided, each SH being able to recognise that his/her specific input has been taken into account proportionately of his/her decisional power; also sensitivity studies must be easy to perform: if some SH wants to check how his/her opinion changes would affect the weight-elicitation results, for example in a Delphi-type enquiry, information about this change should be passed over in a complete and transparent way. Such feature is needed to facilitate the dialogue and convergence between SH, an important aspect in solving ethical dilemmas.

Several well-known techniques with ordinal or rank-based data requiring limited cognitive work have been reviewed for approaching these requirements: in particular the Simos-card

procedure and the surrogate weight techniques [7, 8]. The review showed however that these techniques, although they have inspiring features, also have drawbacks and limited theoretical foundations. It is why a new approach in MCDA is proposed to complement ASCM. The authors propose that SH express qualitative importance judgements on criteria by means of profiles on ordinal Likert scales representing several importance classes. In the following section this new weight-elicitation procedure is introduced. A didactic example is solved, illustrating the fulfilment of the above properties. Conclusions and indication of future work are given in the last section.

WEIGHT ELICITATION IN ASCM

Basic principles

Basic assumptions are that each SH has defined a set of representative criteria, and that their weights can be established independently of possible piloting policies to be ranked by MCDA. On those premises each SH should provide two types of inputs on importance classes and on criterion profiles:

- ✓ a list of naturally ordered importance classes from least to most important: classes define a L-points Likert scale, in short L-pts; each class receives an ordinal number from '1', '2', …, 'L'>1 corresponding to an importance judgment on the criteria to be considered: the lowest class labeled '1' corresponds to a 'least important' judgment, the upper class labeled 'L' corresponds to a 'most important' judgment, of course this most important class is never empty; sometimes it is necessary to add the '0' class of 'not considered' judgments giving a vanishing weight. This is useful when SH propose different sets of criteria, or when transforming different L-pts scales by a proportionality rule to make them comparable (see below); the number of classes is expected to respect the celebrated magic number 'seven plus minus two' [9] indicating cognitive limits about the number of relative importance levels;
- ✓ a singleton profile for each criterion by assigning a value 1 (=100%) to the class corresponding to his/her corresponding importance judgment. A least one criterion must have 100% in the most important class labelled 'L'.

Examples of profiles can be seen in Table 1 defined on three types of scales. Each combined criterion profile is comparable to a probability distribution, i.e., it represents a frequency and it is normalised to unity=100%: the number in each class corresponds to the frequency of opinions in favour of this importance class for the given criterion - 0% or 100% for each individual judgement. It is shown in [10] that the profiles are probability distributions because the importance classes are naturally ordered, so that the feasible class weights, defined up to a positive factor for instance in the [0, 1] interval, are random numbers $\omega_l, l=1,2,\cdots,L$: their probability laws are called order statistics, and they are located in the $L-D$ polytope $\{0<\omega_1<\omega_2<\cdots<\omega_L<1\}$. There is no reason to privilege any set of class-weight values in this polytope, and therefore the class weights are the mean values, indicated by a bar on top of the variables, obeying the order statistics:

$\left(\overline{\omega_l} = l/(L+1), l = 1,2,\cdots,L\right) \sim (1,2,3,\cdots,L)$. As a result the class weights can be set equal to the natural number equivalent to the class labels '1', '2',…, 'L'. The following properties result [10]:

Properties: Given a set of criteria and the associated profiles by criterion, on a common L-pts scale:

(1) The class weights are proportional to the natural numbers $(0,1,2,\cdots,L)$;
(2) The consequence of (1) is that any number of profiles on different scales can be combined by proportionality rules to obtain a global profile on an arbitrary chosen common scale. It is explained in the example below how to proceed;
(3) Each not-normalised criterion weight (NNW), defined up to a multiplicative positive factor, is the mean value of the probability distribution given by the corresponding profile on the class weights $(0,1,2,\cdots,L)$;
(4) Neglecting second-order terms, the normalised criterion weights (NW) are obtained by normalising the NNW. This is only true as a first order approximation because the mean value of the ratio of two random variables is not equal to the ratio of their mean values. Second-order terms can be calculated analytically or by Monte-Carlo simulations and they can be neglected in most cases;

An example for the weight elicitation procedure

A long-term policy in the environment is to be implemented and monitored using ASCM: it implicates all three dimensions in the Golden Triangle. There are three groups of Stakeholders (SH) defining potential policies: the Trade Unions (TU) with a 'Decisional Power' DP=2 and a 4-pts scale, the Government (GV), with DP=2 and a 5-pts scale, and the Environmentalist group (EV), with DP=1 and a 7-pts scale. Three criteria are common to all SH groups and are needed to select a policy: Costs, Emissions (of green-house gases), and Employment. Table 1 shows the three profiles on the corresponding scale for each group: criterion profiles appear to be singletons for each SH group, meaning that all members within a group agree. This last assumption is made for simplicity but without loss of generality, more complicated profiles by groups can be handled in the same way. The upper class on each scale is never empty, i.e., that a least one criterion is 'most important' in each individual profile evaluation. Because of properties (1) and (2) above, transforming each individual profile value from one scale to another different scale may be done by a proportionality rule resulting in a split of each profile value into two values. For example transforming the TU Costs singleton profile '1' in class 3 in the 4-pts scale, meaning a class weight=3, to the EV 7-pts scale gives a doublet profile resulting for the transformed class weight=3x7/4=21/4=5.25 with two values (0.75, 0.25) in the two neighbouring classes (5, 6), as indeed 5x0.75+6x0.25=5.25. The resulting global profiles are given in Table 2, comparing the TU-scale (4-pts) with the EV-scale (7-pts). Not-normalised criterion weights (NNW) are indicated in the corresponding column; they are in the same ratio as the respective numbers of scale points, namely 7/4=1.75, and therefore the normalised (to unity) weights (NW) on the right are the same ones for both scale choices. It is also found with Monte-Carlo simulations that simulated weights (SW) are up to two digits equal to the NW obtained by the simple arithmetic operations.

Table 1: Three reference profiles according to three criteria and three groups of Stakeholders: Trade-Unions (TU) 4-pts Decisional Power DP=2; Government (GV) 5-pts DP=2; Environmentalists (EV) 7-pts DP=1. Alternative individual profiles are indicated in brackets.

Criteria	1	2	3	4	5	6	7
Costs TU (DP=2)	(1)		1				
Emissions	1						
Employment				1			
Costs GV (DP=2)					1		
Emissions			1				
Employment					1		
Costs EV (DP=1)		(1)	1				
Emissions							1
Employment			1				

The resulting normalised weights (NW) are 36%, 24%, 40%. Sensitivity analyses with a spreadsheet or a suitable code, MATLAB® has been used, give immediate results: for example if the TU group brings Costs down from '3' to '1' the new weight values are 29%, 27%, 44%, a 4% increase in favour of the employment weight. If the EV group brings Costs down from '3' to '2', the weights 35%, 25%, 40% result, an increase of 1% only in favour of Emissions. The required visibility features described above are present.

Table 2: Three profiles for two choices of scales 7-pts and 4 pts for the reference case in Table 1. NNW=not-normalised weights in the ratios 7/4, NW= normalized weights, SW=weights obtained by Monte-Carlo simulations (20,000 instances), equal up to two digits.

Criteria	1	2	3	4	5	6	7	NNW	NW=SW
Costs 7-pts	**0**	**0**	**.20**	**0**	**.30**	**.10**	**.40**	**5.50**	**.36**
Emissions	**.10**	**.30**	**0**	**.32**	**.08**	**0**	**.20**	**3.78**	**.24**
Employ.	**0**	**0**	**.20**	**0**	**0**	**0**	**.80**	**6.20**	**.40**
Costs 4-pts	**.06**	**.14**	**.40**	**.40**				**3.14**	**.36**
Emissions	**.40**	**.24**	**.16**	**.20**				**2.16**	**.24**
Employ.	**.06**	**.14**	**0**	**.80**				**3.54**	**.40**

CONCLUSIONS

Implementing ASCM in an efficient way for piloting complex systems requires a simple but accurate and transparent methodology for the weight elicitation for selecting good multiple-criteria policies. The existing techniques often require important cognitive efforts and/or have difficulties in coping with many diverging opinions of multiple stakeholders in most issues addressed by ASCM. Adapting existing rank-based or ordinal techniques [7, 8] has brought forward an approach which is innovative in MCDA: it has a strong theoretical basis [10] and fulfils transparency, simplicity and flexibility requirements. The next planned development is a practical implementation to prove the friendliness and usefulness in real-world systems to be piloted by ASCM. To that aim efficient visualisation instruments

available like in the PROMETHEE methodology [5], and the associated software http://www.promethee-gaia.net/ could further assist the elicitation process of multiple stakeholder groups.

REFERENCES

1. J. Figueira, S. Greco, M. Ehrgott, "Multiple criteria decision analysis: State of the art surveys", Springer Science+Business Media, Boston/Dordrecht, London, 2005.

2. J.P. Brans, C. Macharis, P.L. Kunsch, A. Chevalier, M. Schwaninger, "Combining multicriteria decision aid and system dynamics for the control of socio-economic processes. An iterative real-time procedure", European Journal of Operational Research, 109, 1998, 428-41.

3. J.P. Brans, P.L. Kunsch, and B. Mareschal "Management of the future. A system dynamics and MCDA approach", in, D. Bouyssou, E. Jacquet-Lagrèze, P. Perny, R. Slowinski, D. Vanderpooten, Ph. Vincke (eds.), Special volume dedicated to Professor Bernard Roy, Kuwer Academic Publishers, 2001, 483-502.

4. J.P. Brans, P.L. Kunsch, "Ethics in Operations Research and Sustainable Development", International Transactions In Operational Research (ITOR), 17, 2010, 427-444.

5. J.P. Brans, B. Mareschal, "PROMETHEE methods", in: J. Figueira, S. Greco, M. Ehrgott, "Multiple criteria decision analysis: State of the art surveys", Springer Science+Business Media, Boston/Dordrecht, London, 2005, 163-196.

6. P.L. Kunsch, I. Kavathatzopoulos, F. Rauschmayer, "Modelling complex ethical decision problems with Operations Research", Omega, Special issue on Ethics and Operations Research, 37/6, 2009, 1100-1108.

7. J. Figueira, B. Roy, "Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure", European Journal of Operational Research 139, 2002, 317-326.

8. J. Wang, S. Zionts, "Using Ordinal Data to Estimate Cardinal Values", Journal of Multi-Criteria Decision Analysis, 22, 2015, 185-196.

9. G.A. Miller, "The magic number seven plus or minus two: some limits on our capacity for processing information", Psychological review 13, 1956, 81-97.

10. P.L. Kunsch, A. Ishizaka, "Combining surrogate-weights methods and the Simos procedure in additive multiple-criteria methodologies", in preparation, 2017.

Abel Coll

Abel Coll is a Doctor in Civil Engineering from the Technical University of Catalonia (BarcelonaTech). He is the Head of the Pre and Post-processing department, also known as GiD Group, of the International Center for Numerical Methods in Engineering (CIMNE) at Barcelona, Spain. CIMNE was created in 1987 and specializes in the development and applications of numerical methods and software to find solutions to a variety of problems in engineering and applied sciences.

He begun its research at CIMNE on 2004 focusing on Computational Geometry and Mesh Generation for numerical simulations, and has collaborated in a wide list of European and National research projects related to different scientific and engineering fields (such as Computational Fluid Dynamics, Computational Electromagnetics, Computational Mechanics, Structural analysis or multi-physics problems). His research interest and expertise include preprocessing operations, including computer aided design (CAD) tools, mesh generation and software integration), as well as post-processing operations and visualization of results attached to meshes and coming from numerical simulations. Other research lines of his interest are cloud computing, high performance computing and Big Data for simulation software. Most of this work is part of the development of the GiD suite, a software developed and marketed by the GiD group at CIMNE, for pre and postprocessing data for engineering simulation software (www.gidhome.com).

This talk will present an overview of VELaSSCo project (www.velassco.eu), which is one of the pioneer European projects proposing to merge the Big Data technology with the engineering simulation tools in order to allow the post-processing of huge simulation data (data analytics) and its visualization in an efficient and a more understandable way for the decision making process, what is currently a bottleneck in the simulation pipeline.

Talk: Using Big Data for extreme large scale simulations analysis and visualization.

Numerical simulations are often included in some Decision Support Systems (DSS) when accurate predictions of the physical effects of some events in the domain of study are necessary for the assessment of the best solution. As the phenomena to be studied grows in complexity, or the domain space where the phenomena occurs increases, the related numerical simulation generates greater amounts of data to be analyzed and processed, usually in the range of billions of records.

For these huge simulations it is common to use High Performance Computing (HPC) infrastructures, which typically work on distributed machines with the goal of speed-up the computation times. This generates big amounts of distributed data that need, first, to be post-processed (new data must be generated from them, as iso-surfaces, stream-lines, color maps, etc…) and, second, to be visualized in a reasonably interactive time, almost real-time.

A tool for post-processing and visualize the results of huge numerical simulations in an efficient way is crucial to provide a rapid understanding of the effect of some event in the environment of study for the DSS, and allow a fast integration of this information in the whole decision making process. Furthermore, it is common to use techniques as Artificial Neural Networks (ANN), Model Reduction or optimization techniques based on genetic or Monte-Carlo methods, which also need to run a big number of simulations before the data from the simulation can be reliable. This fact increases even more the data from simulations that need to be processed before feeding the DSS.

VELaSSCo project, Visual Analysis for Extremely Large-Scale Scientific Computing (www.velassco.eu) is a FP7 research project led by CIMNE, funded by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 619439, which has developed a post-processing and visualization platform for huge distributed simulation data using Big Data technologies and architectures. This talk will give an overview of VELaSSCo project and its potential for integrating huge numerical simulations in the DSS.

ICDSST 2017 Full Papers – Abstracts

Business process modelling and visualisation to support e-government decision making: Business/IS alignment

Sulaiman Alfadhel[1], Shaofeng Liu[1], Festus O. Oderanti[2],

[1] Plymouth Graduate School of Management, Plymouth University, UK
[2] Hertfordshire Business School, University of Hertfordshire, UK

sulaiman.alfadhel@plymouth.ac.uk; shaofeng.liu@plymouth.ac.uk;
F.Oderanti@Herts.ac.uk

Abstract. Alignment between business and information systems plays a vital role in the formation of dependent relationships between different departments in a government organization and the process of alignment can be improved by developing an information system (IS) according to the stakeholders' expectations. However, establishing strong alignment in the context of the eGovernment environment can be difficult. It is widely accepted that business processes in the government environment plays a pivotal role in capturing the details of IS requirements. This paper presents a method of business process modelling through UML which can help to visualise and capture the IS requirements for the system development. A series of UML models have been developed and discussed. A case study on patient visits to a healthcare clinic in the context of eGovernment has been used to validate the models.

Keywords: IS requirements, process modelling and visualisation, goal modelling, UML, requirements elicitation.

Value of Visual Analytics to South African Businesses

Wisaal Behardien and Mike Hart

Department of Information Systems, University of Cape Town, South Africa
wisaalbehardien@gmail.com, mike.hart@uct.ac.za

Abstract. There is limited literature on the value that visual analytics provides for businesses, and its broad use in organisations. This research provides some understanding of how South African businesses are using visual analytics in their day to day operations, and the value derived from employing it. The study was interpretive, exploratory and descriptive, producing both quantitative and qualitative data. Individuals within organisations making use of visual analytics completed an online survey, and interviews were conducted with informed business, IT and BI stakeholders. Results were compared with those from an international survey, and thematic analysis highlighted four main themes: usage, value, challenges and technology. Most respondents noted the high added value obtained from visual analytics versus tables of numbers. The research also identified a set of good practices for organisations to employ when embarking on a visual analytics strategy and suggested ways of mitigating potential challenges.

Keywords: visual analytics, visualisation, value, benefits, challenges, business intelligence, data, change management

Conceiving Hybrid What-If Scenarios Based on Usage Preferences

Mariana Carvalho, and Orlando Belo

ALGORITMI R&D Center
Department of Informatics, School of Engineering, University of Minho
4710-057 Braga, Portugal

Abstract. Nowadays, enterprise managers involved with decision-making processes struggle with numerous problems related to market position or business reputation of their companies. Owning the right and high quality set of information is a crucial factor for developing business activities and gaining competitive advantages on business arenas. However, today retrieving information is not enough anymore. The possibility to simulate hypothetical scenarios without harming the business using What-If analysis tools and to retrieve highly refined information is an interesting way for achieving such business advantages. In a previous work, we introduced a hybridization model that combines What-If analysis and OLAP usage preferences, which helps filter the information and meet the users' needs and business requirements without losing data quality. The main advantage is to provide the user with a way to overcome the difficulties that arise when dealing with the conventional What-If analysis scenario process. In this paper, we show an application of this methodology using a sample database, and compare the results of a conventional What-if process and our methodology. We designed and developed a specific piece of software, which aims to discover the best recommendations for What-If analysis scenarios' parameters using OLAP usage preferences, which incorporates user experience in the definition and analysis of a target decision-making scenario.

Keywords: What-If Analysis, On-Line Analytical Processing, Usage Preferences, Analysis Systems Specification, Multidimensional Databases.

Incorporating Uncertainty into Decision-Making: An Information Visualisation Approach

Mohammad Daradkeh and Bilal Abul-Huda

Department of Management Information Systems, Faculty of Information Technology and Computer sciences, Yarmouk University, Irbid, Jordan
{mdaradkeh, abul-huda}@yu.edu.jo

Abstract. Incorporating uncertainty into the decision-making process and exposing its effects are crucial for making informed decisions and maximizing the benefits attained from such decisions. Yet, the explicit incorporation of uncertainty into decision-making poses significant cognitive challenges. The decision-maker could be overloaded, and thus may not effectively take the advantages of the uncertainty information. In this paper, we present an information visualisation approach, called RiDeViz, to facilitate the incorporation of uncertainty into decision-making. The main intention of RiDeViz is to enable the decision-maker to explore and analyse the uncertainty and its effects at different levels of detail. It is also intended to enable the decision-maker to explore cause and effect relationships and experiment with multiple "what-if" scenarios. We demonstrate the utility of RiDeViz through an application example of a financial decision-making scenario.

Keywords: Decision support, information visualisation, interaction design, risk, uncertainty, what-if analysis.

Process Analytics Through Event Databases: Potentials for Visualizations and Process Mining

Pavlos Delias[1] and Ioannis Kazanidis[1]

Eastern Macedonia and Thrace Institute of Technology, Kavala 65110, Greece,
pdelias@teiemt.gr

Abstract. Events, routinely broadcasted by news media all over the world, are captured and get recorded to event databases in standardized formats. This wealth of information can be aggregated and get visualized with several ways, to result in alluring illustrations. However, existing aggregation techniques tend to consider that events are fragmentary, or that they are part of a strictly sequential chain. Nevertheless, events' occurrences may appear with varying structures (i.e., others than sequence), reflecting elements of a larger, implicit process. In this work, we propose several transformation templates to a enable a process perspective for raw event data. The basic idea is to transform event databases into a format suitable for process mining (aka event log) to enable the rich toolbox of process mining tools. We present our approach through the illustrative example of the events that happened in Greece during the referendum period (summer 2015).

Keywords: Event Data, Process Mining, Event Analytics

Multicriteria decision making for healthcare facilities location with visualization based on FITradeoff method

Marta Dell'Ovo[1], Eduarda Asfora Frej[2], Alessandra Oppio[3], Stefano Capolongo[1], Danielle Costa Morais[2] and Adiel Teixeira de Almeida[2]

[1] Dept. ABC, Politecnico di Milano, via Bonardi 9, 20133 Milano, Italy
[2] Federal University of Pernambuco – UFPE, Center for Decision Systems and Information Development – CDSID, Cx. Postal 7462, Recife-PE, 50.630-970, Brazil
[3] Dept. DAStU, Politecnico di Milano, via Bonardi 3, 20133 Milano, Italy
marta.dellovo@polimi.it

Abstract. This paper proposes an application of the Flexible Interactive Tradeoff (FITradeoff) method for siting healthcare facilities. The selection of the location of complex facilities, as hospitals, can be considered as a multidimensional decision problem for the several issues to be taken into account and, moreover, for the variety of stakeholders that should be involved. The case study under investigation is the location of "La Città della Salute", a new large healthcare facility in Lombardy Region (Italy). Starting from a cross disciplinary literature review, a multidimensional evaluation framework has been defined and applied to the case study by considering the point of view of one Decision Maker (DM). The application shows that a smaller effort is required from the DM using the FITradeoff method.

Keywords: Healthcare Facility Location, MCDM, FITradeoff, decision-making, visualization

A semantics extraction framework for decision support in context-specific social web networks

Manuela Freire[1,3], Francisco Antunes[1,2] and João Paulo Costa[1,3]

[1] INESCC – Computer and Systems Engineering Institute of Coimbra, Portugal
[2] Department of Management and Economics, Beira Interior University, Portugal
[3] Faculty of Economics, Coimbra University, Portugal

Abstract. We are now part of a networked society, characterized by the intensive use and dependence of information systems that deals with communication and information, to support decision-making. It is thus clear that organizations, in order to interact effectively with their customers, need to manage their communication activities at the level of online channels. Monitoring these communications can contribute to obtain decision support insights, reduce costs, optimize processes, etc. In this work, we semantically studied the discursive exchanges of a Facebook group created by a strawberries' seller, in order to predict, through Social Network Analysis (SNA) and semantic analysis of the *posts*, the quantities to be ordered by customers. The obtained results show that the unstructured data of the Web's speech can be used to support the decision through SNA.

Keywords: Social Network Analysis, Decision Support, Web Discourse.

A visual decision support system for helping physicians to make a decision on new drugs

Jean-Baptiste Lamy[1], Adrien Ugon[1], Hélène Berthelot[1], Madeleine Favre[2]

[1] LIMICS, Université Paris 13, Sorbonne Paris Cité, 93017 Bobigny, France, INSERM UMRS 1142, UPMC Université Paris 6, Sorbonne Universités, Paris, France,
jean-baptiste.lamy@univ-paris13.fr,
WWW home page: http://www.lesfleursdunormal.fr,
[2] Dept. of primary care, Université Paris Descartes, Société de Formation Thérapeutique du Généraliste (SFTG), Paris, France

Summary. When new drugs come onto the market, physicians have to decide whether they will consider the new drug for their future prescriptions. However, there is no absolute "right" decision: it depends on the physician's opinion, practice and patient base. Here, we propose a visual approach for supporting this decision using iconic, interactive and graphical presentation techniques for facilitating the comparison of a new drug with already existent drugs. By comparing the drug properties, the physician is aided in his decision task.
We designed a prototype containing the properties of 4 new drugs and 22 "comparator" drugs. We presented the resulting system to a group of physicians. Preliminary evaluation results showed that this approach allowed physicians to make a decision when they were lacking information about the new drug, and to change their mind if they were overconfident in the new drug.

Key words: Knowledge visualization, Overlapping set visualization, Medical decision support, Drug knowledge

Implementation of an extended fuzzy VIKOR method based on triangular and trapezoidal fuzzy linguistic variables and alternative deffuzification techniques

Nikolaos Ploskas[1⋆], Jason Papathanasiou[2], and Georgios Tsaples[2]

[1] Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh 15213, PA, USA
[2] University of Macedonia, 156 Egnatia Str., Thessaloniki 54006, Greece

Abstract. Many Multi-Criteria Decision Making (MCDM) problems contain information about the criteria and/or the alternatives that is either unquantifiable or incomplete. Fuzzy set theory has been successfully combined with MCDM methods to deal with imprecision. The fuzzy VIKOR method has been successfully applied in such problems. There are many extensions of this method; some of them utilize triangular fuzzy numbers, while others use trapezoidal fuzzy numbers. In addition, there are many defuzzification techniques available that are used in different variants. The use of each one of these techniques can have a substantial impact on the output of the fuzzy VIKOR method. Hence, we extend the fuzzy VIKOR method in order to allow the use of several defuzzification techniques. In addition, we allow the use of both triangular and trapezoidal fuzzy numbers. In this paper, we also present the implementation of a web-based decision support system that incorporates the fuzzy VIKOR method. Finally, an application of the fuzzy VIKOR method on a facility location problem is presented to highlight the key features of the implemented system.

Key words: multiple attribute decision making, VIKOR, fuzzy, decision support system, defuzzification

⋆ Corresponding author: nploskas@andrew.cmu.edu

Automatic traffic enforcement camera operation, based on a business intelligence system

Sher, Mali[1], Shifrin, Guy[2]

[1]Israel Traffic Police, R&D Unit, Beit Dagan, Israel, msher@police.gov.il
[2] Israel Police Technology Department, BI unit, Jerusalem, Israel, geobisystems@gmail.com

Abstract. Since 2012, a new automatic traffic enforcement camera project has been in operation in Israel. Several databases are included in this project, i.e. sensor data, traffic reports, and road accident records. In 2014 a business intelligence system was developed to obtain all the data from the sensors of the new project and to merge them with the existing data to run the project effectively and efficiently. The aim of this paper is to present the process and the configuration of the business intelligence system, and to present the improvements in all measurements. In this paper we demonstrate the importance of a business intelligence system for operating, engineering, researching and managing aspects of a project.

Keywords: Business Intelligence, Automatic traffic enforcement cameras, Traffic police.

Visualization for decision support in FITradeoff method: exploring its evaluation with cognitive neuroscience

Adiel Teixeira de Almeida[1], Lucia Reis Peixoto Roselli[1]

[1]Federal University of Pernambuco – UFPE, Center for Decision Systems and Information Development – CDSID, Cx. Postal 7462, Recife-PE, 50.630-970, Brazil, almeida@cdsid.org.br
web-page: www.cdsid.org.br

Abstract. FITradeoff method uses a flexible and interactive approach for supporting decisions in multicriteria problems in the context of MAVT (Multi-Attribute Value Theory) with partial information. Since the very beginning of the preference elicitation process, a subset of potential optimal alternatives (POA) is selected based on the current partial information provided. Then, the Decision Maker (DM) has the flexibility of interrupting the elicitation process for analyzing the partial result by other means, such as graphical visualization of performance of POA. This flexibility is available in the whole process. Evaluating the visualization confidence for decision support in FITradeoff method is crucial. Furthermore, information for designing of this visualization is relevant. This paper shows how these issues could be approached based on cognitive neuroscience, with particular focus given on eye tracking resources.

Keywords: Evaluating visualization, graphical visualization, FITradeoff MCDM method, cognitive neuroscience, neuroeconomics, Decision Neuroscience, eye tracking.

A tool for energy management and cost assessment of pumps in Waste Water Treatment Plants

Dario Torregrossa[1], Ulrich Leopold[1] F. Hernández-Sancho[2], J. Hansen[3], A. Cornelissen[4], and G. Schutz[4]

[1] Luxembourg Institute of Science and Technology (LIST), 41 Rue du Brill, 4422 Sanem, Luxembourg
dario.torregrossa@list.lu,
[2] Estructura Económica, Universitat de València, Avda dels Tarongers, s/n, 46022 Valencia, Spain,
[3] Université du Luxembourg, 6 rue Richard Coudenhove-Kalergi, L-1359 Luxembourg
[4] RTC4Water s.a.r.l, 9, av. des Hauts-Fournaux, 4362 Esch-Sur-Alzette (Belval), Luxembourg

Abstract. Waste Water Treatment Plants (WWTPs) are generally considered energy intensive. Substantial energy saving potentials have been identified by several authors. Pumps consume around 12% of the overall WWTP energy consumption. In this paper we propose a methodology that uses the sensors commonly installed in WWTPs, such as volume and energy sensors, to perform energy benchmarking on pumps. The relationship between energy efficiency and flow rate is used to detect specific problems, and potential solutions are proposed, taking into consideration economical and environmental criteria (cost of externalities in energy production). The methodology integrates energy benchmarking, data-mining, and economical and environmental assessment. In order to make better informed decisions, plant managers can now perform a multivariate analysis within a very short time, using information generally available in WWTPs.

Keywords: Waste Water Treatment Plant (WWTPs), Energy Benchmarking, Pump performance analysis, Externalities cost

Integrating System Dynamics with exploratory MCDA for robust decision-making

Georgios Tsaples[1]*, Jason Papathanasiou[1], and Nikolaos Ploskas[2]

[1] University of Macedonia, 156 Egnatia Str., Thessaloniki 54006, Greece
[2] Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh 15213, PA, USA

Abstract. The aim of this paper is to propose a process to support decision making, in which System Dynamics is combined with Multi Criteria Decision Aid methods to mitigate the limitations of the two methodologies when used alone and find robust policies. The proposed process is based on Exploratory Modeling and Analysis, a framework that allows the use of multiple methods – under different perceptions, detail, and levels of abstraction – in order to address issues of uncertainty and robustness. A case study is used to illustrate how the process can offer deeper insights and act as a valuable decision support system. Finally, it also demonstrates the potential of Exploratory Modeling and Analysis to deal with uncertainties and identify robust policies.

Keywords: Exploratory Modeling and Analysis, System Dynamics, Multi Criteria Decision Aid, decision support system

* Corresponding author: gtsaples@uom.edu.gr

ICDSST 2017 Short Papers

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

A DSS for constructing the Precise Consistency Consensus Matrix in AHP-Group Decision Making[1]

Juan Aguarón, María Teresa Escobar, José María Moreno-Jiménez, Alberto Turón
Grupo Decisión Multicriterio Zaragoza. Facultad de Economía y Empresa
Universidad de Zaragoza. Gran Vía, 2. 50005 Zaragoza, Spain
aguaron@unizar.es, mescobar@unizar.es, moreno@unizar.es, turon@unizar.es
web-page: http://gdmz.unizar.es/

ABSTRACT

The Precise Consensus Consistency Matrix (PCCM) is a decisional tool developed by the Zaragoza Multicriteria Decision Making Group (GDMZ) for dealing with AHP-Group Decision Making (AHP-GDM). From the initial decision makers' pairwise comparison matrices, a consensus matrix for the group is constructed using the concept of consistency. Each entry of this new consensus matrix belongs to all the Consistency Stability Intervals associated to every decision maker. In this way, it is guaranteed that the modifications made in the initial matrix do not exceed the maximum level of inconsistency allowed.

This paper presents a decision support system (PRIOR-PCCM) that allows the construction of the PCCM in the AHP-GDM context. PRIOR-PCCM incorporates the possibility to consider different weights for the decision makers and includes a module that allows the extension of the initial PCCM in order to achieve the minimum number of non-null entries required to derive the priorities, or to achieve a complete PCCM matrix. The DSS also calculates four indicators, two for consistency and two for compatibility, used for comparing the PCCM with other AHP-GDM tools. One of these indicators is employed to identify the decision maker that has modified his judgements most in order to obtain the consensus matrix. This is illustrated in a case study.

The paper also presents some new visualisation tools that can be employed to obtain a deeper knowledge of the process followed to obtain the PCCM matrix and to exploit the results from a cognitive orientation. These new tools are currently being integrated into the DSS software.

Keywords: DSS, Analytic Hierarchy Process (AHP), Group Decision Making (GDM), Consensus, Consistency, Compatibility.

[1] This research was funded by the Spanish Ministry of Economy and Competitiveness along with FEDER funding (Project ECO2015-66673-R).

INTRODUCTION

Two of the most outstanding features of the Analytic Hierarchy Process (AHP; [1]) are the possibility of evaluating the consistency of the pairwise comparison matrices used in order to capture the preferences of the decision makers (DMs) and the adaptability to the Group Decision Making, an outstanding aspect of the Knowledge Society.

The idea of using the concept of consistency in group decision making was first proposed by the authors in 2002 [2, 3]. Afterwards, it is being widely used in the scientific literature [4-7].

Following this line of research, the authors proposed the PCCM [8, 9], a decisional tool for AHP multiactor decision making which main aim is the construction of a consensus matrix based on the consistency property.

The aim of this current work is to present the DSS (PRIOR-PCCM) designed by the GDMZ for constructing the PCCM and to propose some visualization tools that will help to understand the process followed in the application of the algorithm (construction of the consensus matrix) and to exploit the results that are obtained from a cognitive perspective.

The paper is structured as follows: Background section reminds the basics of the decision-making tool (PCCM) and the algorithm followed for its construction; the following section includes a description of the DSS designed for the construction of the PCCM; the fourth section presents an illustrative numerical example to which the DSS and the visualisation tools were employed; and finally, the last section summarises the main conclusions and outlines lines for future research.

BACKGROUND

Saaty [1] defines the consistency as the *cardinal transitivity* between the judgments. Given a pairwise comparison matrix $A = (a_{ij})$, $i,j = 1,\ldots,n$, A is said to be *consistent* if $\forall i,j,k\ a_{ij} \cdot a_{jk} = a_{ik}$. In order to measure the inconsistency of a matrix, the PCCM employs the *Geometric Consistency Index* (GCI) [10].

Following the line of research that uses the idea of consistency in AHP-GDM, the authors [11] proposed an initial decisional tool, named the *Consensus Consistency Matrix* (CCM) that identified the core of consistency of the group decision using an interval matrix, which may not be complete or connected. In 2014, the same authors further refined this tool and introduced the *Precise Consensus Consistency Matrix* (PCCM) [8], which selects a precise value for each interval judgement in such a way that the quantity of slack that remains free for successive algorithm iterations is the maximum possible. Escobar et al. [9] extended the PCCM to allow the assignment of different weights to the decision makers and to guarantee that the group consensus values were acceptable to the individuals in terms of inconsistency. In the same work, these authors proposed different methods for completing the PCCM matrix if it were incomplete. The improved version of the algorithm for constructing the PCCM can be seen in [9]. Figure 1 shows a flowchart of this algorithm.

A DSS FOR THE CONSTRUCTION OF THE PCCM

The Software

PRIOR-PCCM is a DSS developed by the GDMZ for constructing the PCCM. The software was programmed in Delphi and completes other modules previously incorporated into the PRIOR software [12-14]. PRIOR-PCCM includes several modules for: (i) the

calculation of the consistency stability intervals; (ii) the resolution of the optimisation problem; (iii) the derivation of priorities in the case of incomplete matrices; and (iv) the evaluation of consistency and compatibility.

The main window of the DSS (see Figure 2) offers options to introduce a new problem ('Parameters' and 'Judgments'); to read a previously fixed problem ('Load Data'); to calculate the individual priorities ('Priorities'); to calculate the Initial Stability Intervals ('Stability Intervals'); to calculate the AIJ and PCCM matrices ('AIJ Matrix' and Consensus(PCCM)' respectively) and the indicators used to compare both AHP-GDM tools ('CVNs', 'GCOMPIs' and 'PVNs'). It is also possible to edit the Consensus Matrix ('Edit Consensus Matrix'). This option is particularly helpful if the resulting PCCM is incomplete.

When the PCCM matrix is obtained, the corresponding priority vector and GCI are also calculated (see Figure 3). If the resulting matrix (PCCM) is not complete, the DSS allows us to edit the consensus matrix and add new judgements (for different options, see [9]). A minimum of n-1 connected entries are needed to be able to obtain the corresponding priority vector, which can be achieved from this same window. These user-defined matrices can be saved in order to retrieve them when needed.

Once the PCCM has been obtained, the DSS provides the calculation of other indicators (CVN, GCOMPI, PVN) that measure the compatibility of the individual positions with the group one, and compares the two AHP-GDM procedures currently implemented in the software, the AIJ and the PCCM (see Figure 4). All the windows that show some results can be copied for use outside the DSS (Excel, etc.).

Visualisation Tools

A visualisation tool based on Multidimensional Scaling has been designed and developed with R in order to show the process followed to obtain the PCCM. This tool provides some graphics of the evolution of the priorities of the different decision makers throughout the iterations of the algorithm followed to construct the consensus matrix. An example of this visualisation tool can be seen in Figures 5 and 6.

In addition, another visualisation tool is being designed to present graphically the information about the evolution of the indicators calculated by the DSS. Figures 7 and 8 are examples of this tool for the GCI and the GCOMPI. These new tools are being integrated into the DSS software. They will improve the presentation of the final results, allowing the exploitation of the results from a cognitive orientation.

ILLUSTRATIVE EXAMPLE

We have applied the DSS to the case study selected by [8, 9]. The problem consists of five alternatives and three decision makers. Table 1 shows the initial judgement matrices. Different weights were associated to the decision makers (DM1: 5; DM2: 4; and DM3: 2). By using the DSS, the PCCM matrix is obtained. The corresponding priority vector and the value of the GCI (0.023) are also shown (see Figure 3). As can be observed, in this particular case, the PCCM is complete. Figure 5 shows the evolution of the relative positions of the DMs and the alternatives throughout the iterations of the algorithm. It can be appreciated that DMs have quite similar opinions and there are no big modifications in their relative positions. Figure 6 shows a zoom of this graph showing the evolution of DMs with more detail. Figure 7 presents the evolution of the GCI for the three DMs throughout the iterations. It can be seen that although the values of the GCI for some intermediate iterations have increased, the final GCI is notably better than the initial individual ones.

Table 1: Pairwise comparison matrices for the three decision makers

DM1	A	B	C	D	E
A	1	3	5	8	6
B		1	3	5	4
C			1	3	2
D				1	1/3
E					1

DM2	A	B	C	D	E
A	1	3	7	9	5
B		1	3	7	1
C			1	5	1/5
D				1	1/5
E					1

DM3	A	B	C	D	E
A	1	5	7	7	5
B		1	1	5	1
C			1	5	1/3
D				1	1/5
E					1

In order to identify the decision maker who has modified his judgments most, we propose to employ the *Geometric Compatibility Index* (GCOMPI) that compares the initial judgements of a given decision maker with the priorities associated to the final PCCM matrix (see definition in [8, 9]). Figure 4 shows the values of this indicator calculated by the DSS for the example. The values of the GCOMPI for each of the three decision makers are given in the first column. The second column contains the weights associated to each decision maker. The third column calculates the contribution of each decision maker to the value of the GCOMPI for the group ($GCOMPI^{(G)} = 0.529$) and the last column contains this contribution in relative terms. Figure 8 shows a graphic presentation of all these values. It can be seen that although decision-maker 3 has the greatest value of the GCOMPI, the contribution of this decision-maker to the group GCOMPI is the least, as its weight is the smallest. The second decision-maker made the greatest contribution to the total amount of the group GCOMPI.

CONCLUSIONS

This paper presents the PRIOR-PCCM, a DSS that implements the PCCM. This is a decisional tool for AHP-GDM that, based on the concept of consistency, provides a consensus matrix for the group. The paper also presents some new visualisation tools that will favour the exploitation of the model from a cognitive orientation. These new tools, as well as the implementation of other AHP-GDM consensus methods, are currently being integrated into the DSS software.

REFERENCES

1. T.L. Saaty, “Multicriteria Decision Making: The Analytic Hierarchy Process”, Mc Graw-Hill, New York, 1980.

2. J.M. Moreno-Jiménez, J. Aguarón, M.T. Escobar, “Decisional Tools for Consensus Building in AHP-Group Decision Making”, 12th Mini Euro Conference, Brussels, 2002.

3. J.M. Moreno-Jiménez, J. Aguarón, A. Raluy, A. Turón, “A Spreadsheet Module for consistent consensus building in AHP-Group Decision Making”, Group Decision and Negotiation, 14(2), 89-108, 2005.

4. Y. Dong, G. Zhang, W.H. Hong, Y. Xu, “Consensus models for AHP group decision making under row geometric mean prioritization method”, Decision Support Systems, 49, 281-289, 2010.

5. Z.B. Wu, J.P. Xu, “A consistency and consensus based decision support model for group decision making with multiplicative preference relations”, Decision Support Systems, 52, 757-767, 2012.

6. G.Q. Zhang, Y.C. Dong, Y.F. Xu, “Linear optimization modelling of consistency issues in group decision making based on fuzzy preference relations”, Expert Systems with Applications, 39, 2415-2420, 2012.

7. G.Q. Zhang, Y.C. Dong, Y.F. Xu, “Consistency and consensus measures for linguistic preference relations based on distribution assessments”, Information Fusion, 17, 46-55, 2014.

8. J. Aguarón, M.T. Escobar, J.M. Moreno-Jiménez, “The precise consistency consensus matrix in a local AHP-group decision making context”, Annals of Operations Research, 245 (1), 245-259, 2016.

9. M.T. Escobar, J. Aguarón, J.M. Moreno-Jiménez, “Some Extensions of the Precise Consistency Consensus Matrix”, Decision Support Systems, 74, 67-77, 2015.

10. J. Aguarón, J.M. Moreno-Jiménez, “The geometric consistency index: approximate thresholds”, European Journal of Operational Research, 147(1), 137-145, 2003.

11. J.M. Moreno-Jiménez, J. Aguarón, M.T. Escobar, “The Core of Consistency in AHP-Group Decision Making”, Group Decision and Negotiation, 17(3), 249-265, 2008.

12. J. Aguarón, M.T. Escobar, J.M. Moreno-Jiménez, A. Turón, “PRIOR-WEB: A discrete multicriteria prioritization tool for the global economy context”, in A. Respicio et al. (eds.), “Bridging the Socio-technical Gap in Decision Support Systems”, IOS Press, 249-258, Amsterdam, 2010.

13. A. Turón, J. Aguarón, J, Cardeñosa, M.T. Escobar, J.M. Moreno-Jiménez, J. Ruiz-Mas, A. Toncovich, “PRIOR-W&K: A Collaborative Tool for Decision Making in the Knowledge Society”, CCIS, 278, 155-161. Springer Heildelberg, Berlin, 2013.

14. A. Turón, J. Aguarón, M.T. Escobar, C. Gallardo, J.M. Moreno-Jiménez, J.L. Salazar, “PRIOR-WK&E: Social Software for Policy Making in the Knowledge Society”, CCIS, 111, 139-149. Springer Heildelberg, Berlin, 2010.

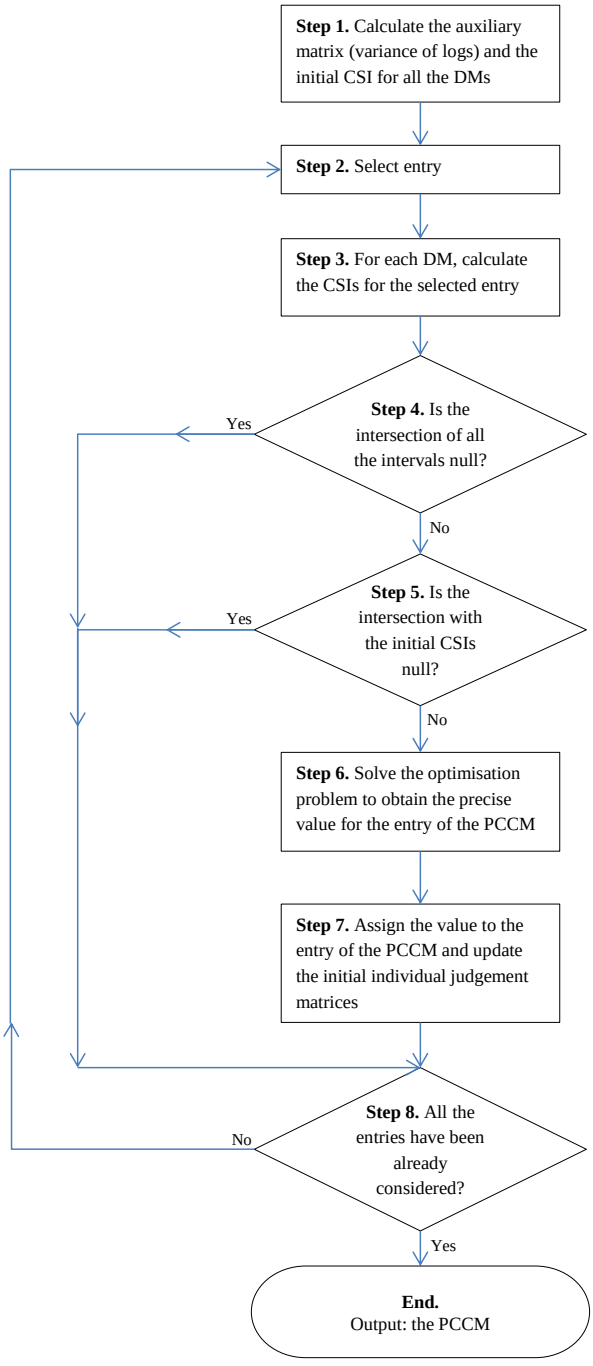


Figure 1: The PCCM algorithm

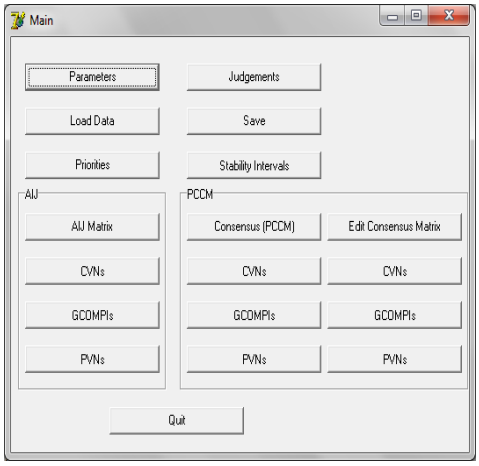


Figure 2: Main window of PRIOR-PCCM

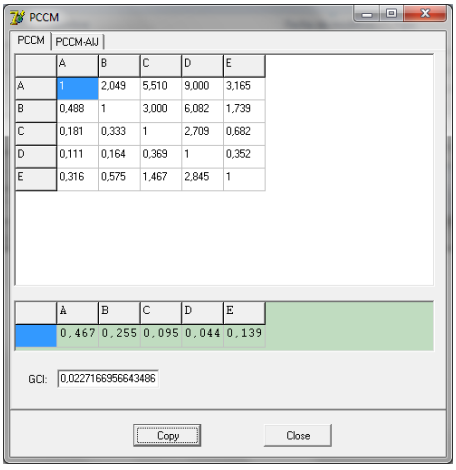


Figure 3: PCCM, priority vector and GCI

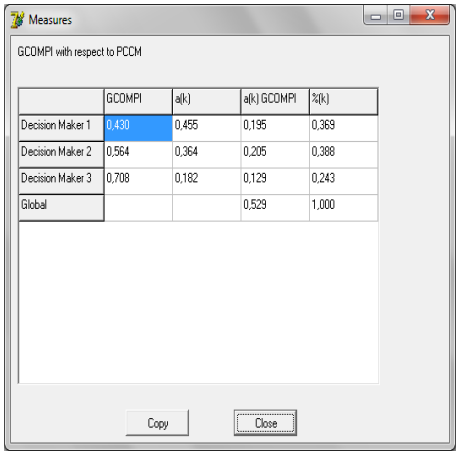


Figure 4: Results for the indicator GCOMPI

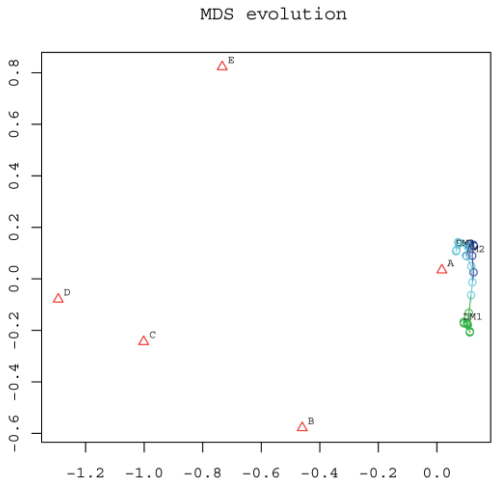


Figure 5: Evolution of the relative positions of the DMs and the alternatives

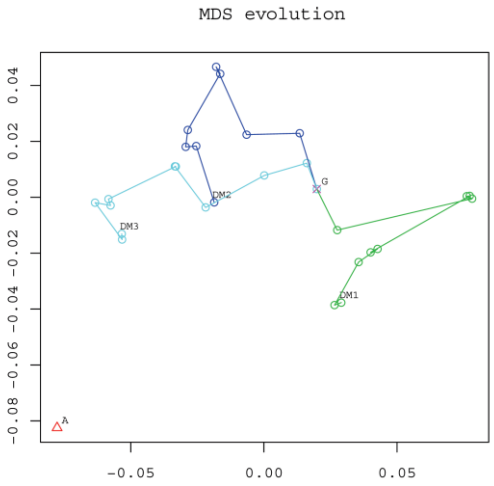


Figure 6: Zoom on the evolution of the relative positions of the DMs

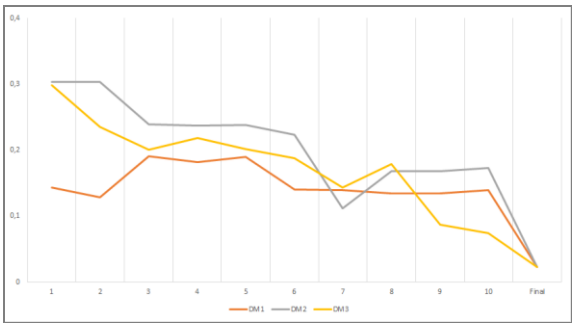


Figure 7: Evolution of the GCI

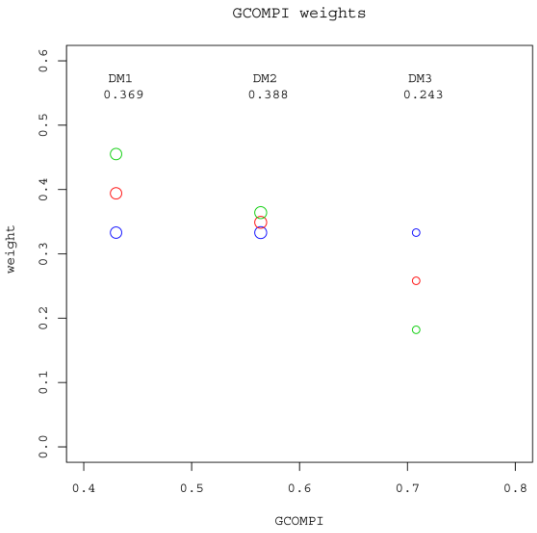


Figure 8: Contribution of the DMs to the GCOMPI

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

Towards a New collaborative Decision process Using Fuzzy Measures and Videoconferencing for disaster management in Algeria

Latifa BOUDJELLALI, Noria TAGHEZOUT
Sid Ahmed EL ANDALOUSSI
Laboratory LIO, Ahmed Ben Bella Oran 1 university, Oran, Algeria
Boudjellali2latifa@gmail.com, taghezout.nora@gmail.com, s-elandaloussi@hotmail.com

ABSTRACT

Algeria is exposed to many kinds of disaster, vulnerable to natural phenomena, such as earthquakes and floods, and to man-made vulnerabilities like road accidents.

Collaborative is the key to the success of the organizational plan that is initially established by relief because many parties are involved.

For this reason, we propose a decision support platform for disaster management using case-based reasoning (CBR) and fuzzy reasoning. This platform offers innovative collaboration through videoconferencing communication between the leaders of intervention services (modules) who are in charge of a mission and the means necessary to relay the rescue operations in a very short time.

Under the authority of the 'Wali'(the president of the City), the relief organization plan is drawn up by the Directorate of Civil Protection, together with the heads of the modules (Security and Public order, Medical evacuation and hygiene order, Liaison and Telecommunication, Information, Public works, Energy, Hydraulic, Transportation, Provisional recessing, Supply of food and Relief, Materials and equipments, Expertise and advice, Evaluation and Balance sheet).

Our approach promotes exchanges among all participants specially the domain experts. The case based reasoning is chosen in order to use past experiences while integrating expert's beliefs. An ontology that represented the most important knowledge of disasters has been developed while integrating experts' beliefs. The particularity of this approach is that it proposes a new adaptation technique which integrates fuzzy measures. Thanks to videoconferencing tools, experts can intervene at any time to give their opinions during the 'revise' phase. The system is being tested to analyze the impact of collaboration tools on disaster management in Algeria, particularly in the western part of the country.

Keywords: disaster management in Algeria, Collaboration, Case-Based Reasoning (CBR), fuzzy measures, and videoconferencing.

INTRODUCTION

The natural disasters are caused by natural or technological reasons that could affect life, society and the environment. Potential consequences of disasters involve overwhelming economic losses, large affected populations and serious environmental damages [1].

Proper preparation, response and recovery are required in order to cope with and survive large-scale disasters; unfortunately, the institutions that are responsible for delivering emergency response services form a heterogeneous set that often under-performs are due to lack of proper interoperation and collaboration [2]. Collaboration is a necessary foundation for dealing with both natural and technological hazards as well as disasters and the consequences of terrorism [3].

Collaboration and videoconferencing systems have become a very important application in the Internet. The paper adopts an interdisciplinary approach to disaster management collaboration drawing upon knowledge accumulated by state-of-the-art research in case-based reasoning (CBR) and videoconferencing systems.

Major contribution

In recent years, social media emerged as a potential resource to improve the management of crisis situations such as disasters triggered by natural hazards. Although there is a growing research body concerned with the analysis of the usage of social media during disasters, most previous work has concentrated on using social media as a stand-alone information source, whereas its combination with other information sources holds a still underexplored potential [4]. For this reason we propose to elaborate a collaborative environment that contains a social network in order to collect information in the event of a disaster.

The particularity of our approach is in the use of the case-based reasoning. This process follows five steps (Elaborate, Retrieve, Reuse, Revise and Retain) including domain ontology for the elaboration of the new case, fuzzy logic for Reuse step, and video conferencing in revision because a video communication tool must always be available (join these contact (Experts) should be simple and fast, as well as investments of time and limited money).

Our approach

Case Based Reasoning (CBR) is an Artificial Intelligence (AI) Technique for solving problems by using past experience instances. The centrality of this concept lies in using past solved problems in similar situations to approximate past solutions for such problem situations.

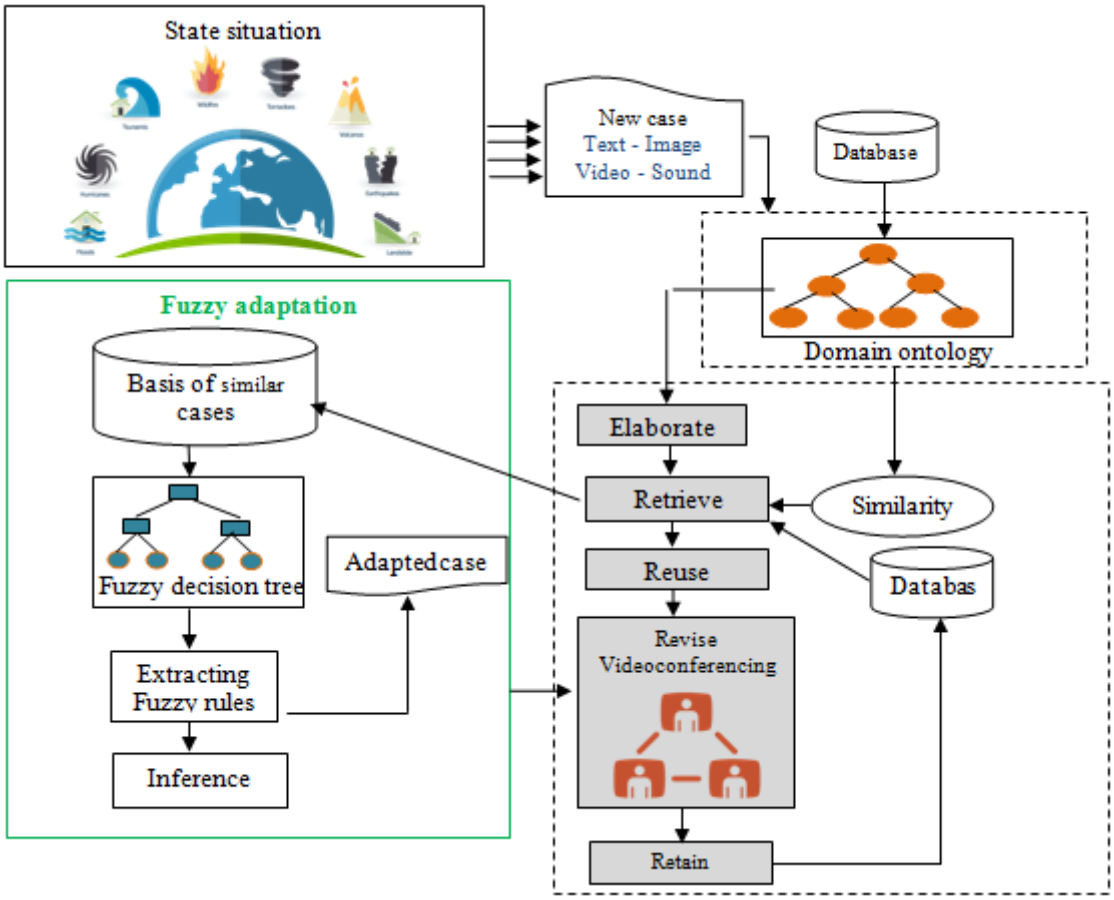


Figure 1: General scheme of the proposed approach

Case elaboration

A new case is described by a set of descriptors; a descriptor of a pair is defined by "d = (a, v) where ‘a’ is an attribute and ‘v’ 'is a value associated with ‘a’ in this case, to constitute this case the user often acquires non-elaborated case where some information are unknown or missing, it’s important to improve the description of the problem; So we propose to improve the development of a case by constructing the domain ontology.

Information retrieval

Information retrieval about similar cases requires an efficient comparison between the descriptors for the similarity evaluation. Considering the wide range of available semantic similarity approaches, our work focused on the taxonomical exploitation of the ontology.

Reuse case (fuzzy adaptation)

Information Fuzzy logic allows reasoning not on numeric variables, but on linguistic ones, ie, on qualitative variables (large, small, medium, far, near, safe, etc.). Reasoning on these linguistic variables will allow us to use the knowledge in natural language.

In inductive learning, most models use classical logic representation and processing of ambiguities and uncertainties throughout the induction process. Table 1 illustrates the CBR that is guided by fuzzy logic in some works.

Table 1: Elements of comparison between some works (CBR & Fuzzy logic).

The references	relevant phase	Technique used
[05] [06]	Information retrieval	Incorporating fuzzy logic in the CBR process for a model prediction of the previous similar auctions and more useful to a current case.
[07]	Information retrieval	Introducing fuzzy logic in calculating the similarity between two cases statewide "Quite similar", "somewhat similar" or "very similar".
Our approach	Fuzzy adaptation	Introduces the fundamentals of fuzzy logic for knowledge representation, and the fuzzy decision tree for the management of knowledge.

We use the notions of fuzzy logic that are made by author in [08] in our approach because the fuzzy logic allows representation and adequate treatment of the inaccuracies and uncertainties priori similar to those used by humans, and this can be done through its various concepts such the fuzzy sets and fuzzy reasoning. So we have opted for the use of fuzzy logic in the induction process by utilizing fuzzy decision tree to integrate it in the two characteristics of intelligence: knowledge management inaccuracies and uncertainty management in the provided data.

Revise

Realizing the fourth stage of the CBR must be made with the contribution of experts to review the obtained case.

The goal of real-time telecommunication media is to collapse the space between geographically dispersed groups and create the illusion that people are together, when in fact they are not. For this reason we propose the Videoconferencing technology in this step.

Retrain

After the solution has been successfully adapted to the target problem, store the resulting experience as a new case in the case base.

EMERGENCY ORGANIZATION PLAN

The Government of Algeria adopted, in 1985 [09], a national plan for the prevention of major natural and technological hazards, mainly involving the organization of interventions and disaster relief.

The plan for the organization of interventions and relief in the state is made up of 14 intervention modules.

Under the authority of the Wali, it is elaborated by the Directorate of Civil Protection, together with the heads of modules. We propose collaboration between the latter by videoconferencing. Figure 2 depicts the major steps that are followed to execute the emergency organization plan.

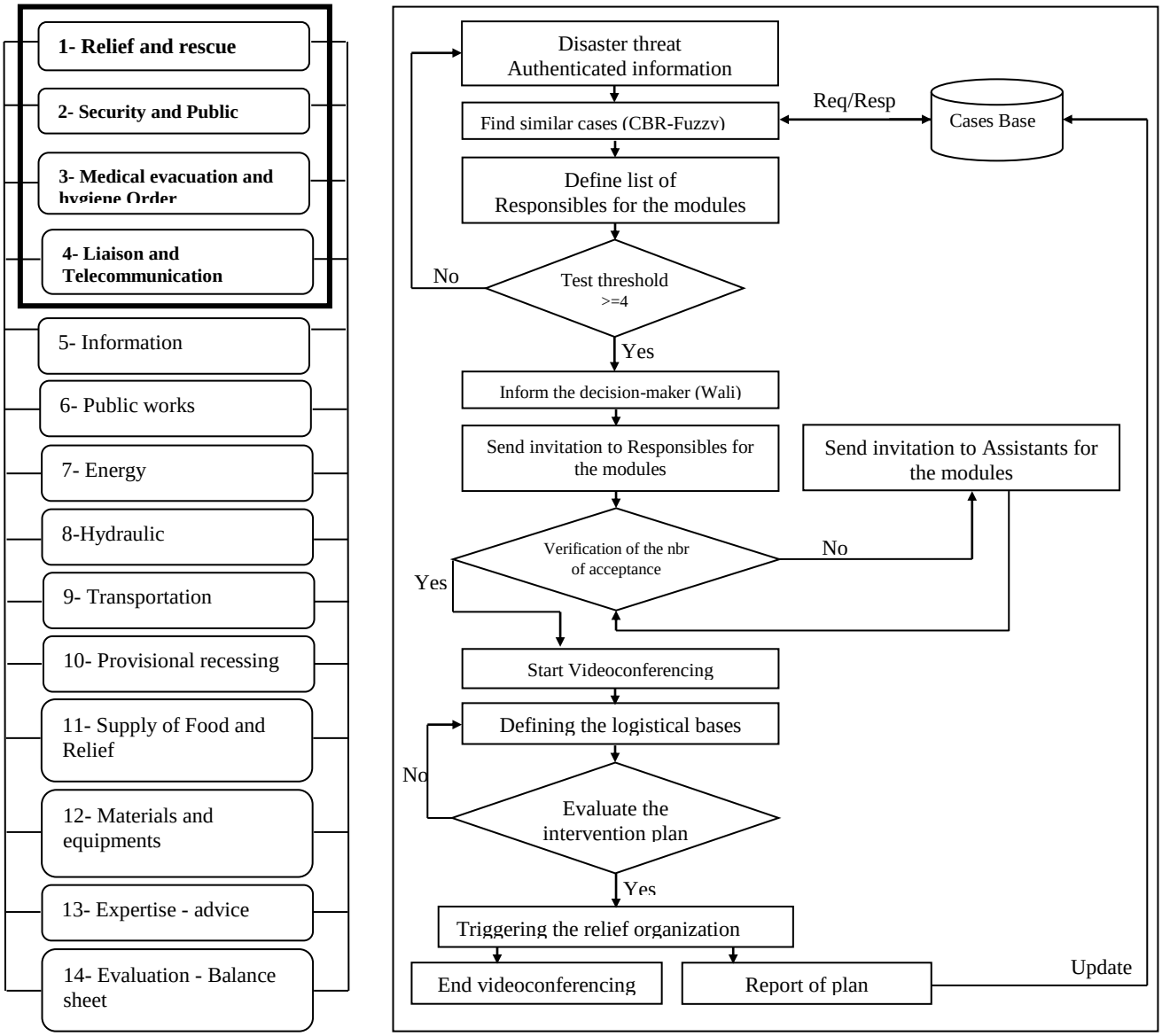
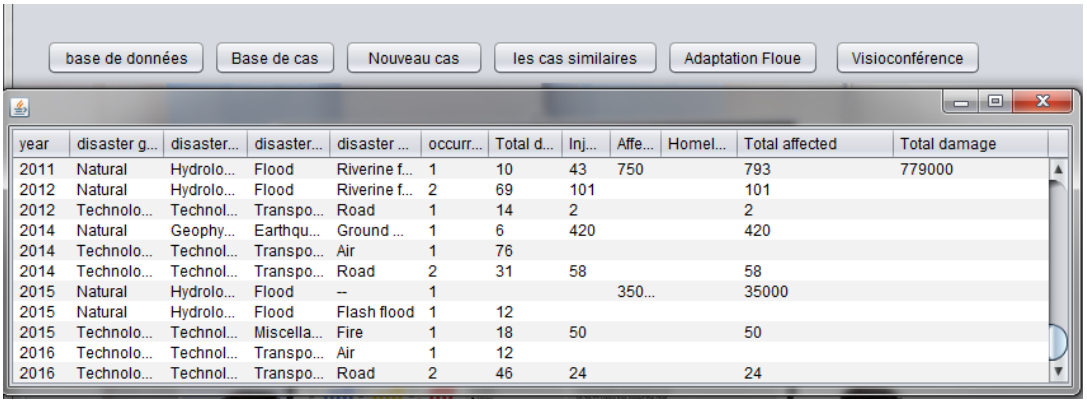


Figure 2: Videoconferencing for emergency organization plan

In order to validate our approach we apply it on the basis EM-DAT[10], available on the official website EM-DAT website. EM-DAT contains essential data on the occurrence and effects of over 22,000 mass disasters in the world from 1900 to the Present day. The most relevant cases (107 disaster cases in Algeria) are presented in Figure3.



year	disaster g...	disaster...	disaster...	disaster ...	occurr...	Total d...	Inj...	Affe...	Homel...	Total affected	Total damage
2011	Natural	Hydrolo...	Flood	Riverine f...	1	10	43	750		793	779000
2012	Natural	Hydrolo...	Flood	Riverine f...	2	69	101			101	
2012	Technolo...	Technol...	Transpo...	Road	1	14	2			2	
2014	Natural	Geophy...	Earthqu...	Ground ...	1	6	420			420	
2014	Technolo...	Technol...	Transpo...	Air	1	76					
2014	Technolo...	Technol...	Transpo...	Road	2	31	58			58	
2015	Natural	Hydrolo...	Flood	--	1			350...		35000	
2015	Natural	Hydrolo...	Flood	Flash flood	1	12					
2015	Technolo...	Technol...	Miscella...	Fire	1	18	50			50	
2016	Technolo...	Technol...	Transpo...	Air	1	12					
2016	Technolo...	Technol...	Transpo...	Road	2	46	24			24	

Figure 3: disasters database in Algeria

CONCLUSIONS

Our study presents the relief organization plan in Algeria; this plan is made up of 14 intervention modules. Each module has a logistic base (set of missions and means).

In this paper, we describe an approach of CBR that is combined with the disasters ontology and concepts of fuzzy logic. The proposed solution provides expert knowledge appropriate. One way to reduce costs is to hold a meeting remotely by a videoconferencing system.

Our collaborative platform supports the leaders of intervention services and experts in the disaster planning process and relies on videoconferencing to analyze situations and formulate decisions to trigger the plan in real time. The system is being tested to analyze the impact of collaboration tools on disaster management in Algeria, particularly in the western part of the country.

REFERENCES

1. G. Galindo and R. Batta, "Review of recent developments in OR/MS research in disaster operations management", European Journal of Operational Research, vol. 230, no. 2, pp. 201-211, 2013.

2. O. Noran, "Collaborative disaster management: An interdisciplinary approach", Computers in Industry, vol. 65, no. 6, pp. 1032-1040, 2014.

3. W. Waugh and G. Streib, "Collaboration and Leadership for Effective Emergency Management", Public Administration Review, vol. 66, no. 1, pp. 131-140, 2006.

4. J. Albuquerque, B. Herfort, A. Brenning and A. Zipf, "A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management", International Journal of Geographical Information Science, vol. 29, no. 4, pp. 667-689, 2015.

5. P. Chang, C. Liu and R. Lai, "A fuzzy case-based reasoning model for sales forecasting in print circuit board industries", Expert Systems with Applications, vol. 34, no. 3, pp. 2049-2058, 2008.

6. Mekroud, Noureddine, and Abdelouahab Moussaoui. "Approche Hybride pour le Diagnostic Industriel basée sur le RàPC et le Datamining: Utilisation de la Plateforme JCOLIBRI 2.1." CIIA. 2009.

7. Devèze, Benjamin, and Matthieu FOUQUIN. "Case-Based Reasoning." Rapport des études en spécialité SCIA, école de l'ingénieur EPTA, France, 2004.

8. L. Zadeh, "A computational approach to fuzzy quantifiers in natural languages", Computers & Mathematics with Applications, vol. 9, no. 1, pp. 149-184, 1983.

9. Bellala, 'Informations depuis la direction générale de protection civile, ORAN', technical Report, 2013.

10. D. Guha-Sapir, R. Below, Ph. Hoyois - EM-DAT: The CRED/OFDA International Disaster Database – www.emdat.be – Université Catholique de Louvain – Brussels – Belgium.

11. L. Boudjellali and N.Taghezout, Collaborative environment based on videoconferencing for disaster management, Pattern Analysis and Intelligent Systems (PAIS'2016), Khenchela, Algeria, 2016.

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

Implementing Voting Tools in GRUS

Adama Coulibaly[1,2], Pascale Zaraté[1], Guy Camilleri[1],

[1] University of Toulouse / IRIT, Toulouse, France
[2] University of Bamako / USTTB-FST, Bamako, Mali
{Adama.Coulibaly, Pascale.Zarate, Guy.Camilleri}@irit.fr

ABSTRACT

The aim of this work is to develop a recommendation system to help to choose a voting procedure. For this purpose, we have to implement these voting procedures in a GDSS (Group Decision Support Systems). Therefore, it involves to review the literature on voting procedures and voting tools in order to know their functionality. We will use a practical example to show the level of implementation of the platform in its current state. Finally, we will invoke the perspectives of future work to continue to improve the voting tools of GRUS (GRoUp Support).

Keywords: Decision making, voting procedures, voting tools, GDSS, GRUS.

INTRODUCTION AND BACKGROUND

Decision-making and its execution are the fundamental goals of any organization and their management. Indeed, any organization depends on the nature of the decisions that are made within it by decision-makers, whether individual or collective [1].

However, a collective decision often creates conflict situations due to differences in the views and interests of decision-makers about the same set of objects, hence using decision support systems is needed. Making a decision is choosing from a set of alternatives likely to solve a problem in a given context [2]. To achieve this, the problem of collective multicriteria decision-making often comes down to an aggregation problem that is characterized by the definition of a set of criteria to evaluate the performance of the alternatives according to each criterion considered as relevant by all decision makers.

Group Decision Support Systems (GDSS) are developed to help decision makers and are most often based on computer platforms that provide decision-makers with a formal framework for reflection, and investigative skills to express the preferences and parameters of each, to evaluate them, and to provide the relevant elements for the decision-making process.

This type of system consists in offering tools of software for group decision [3]. A particular actor stands out in the process of group decision making. This is the facilitator. This actor's role is to help the group making a decision. This assistance can be defined not only on the technical level, but also on the content or the decision-making process. Few software packages are currently being developed to assist the facilitator in assisting the decision-making process. A Group DSS has been developed at IRIT called GRUS (GRoUp Support) [4]. Indeed, the main

objective of this work is to develop a complementary tool implementing different voting procedures [5] but above all allowing the facilitator to choose the most appropriate procedure depending on the type of the working context of the group through a recommendation mechanism. The working context of the group is defined by the size of the group, the date of delivery of the decision, the working mode of the group (synchronous or asynchronous / distributed or not).

To achieve these objectives, we must make a state of the art on the different existing voting procedures, to understand their rules of operation. In addition, we also have to know what are the existing tools implementing the voting procedures.

RELATED WORK

Voting procedures

Voting is considered as an individual mode of expression allowing a group of people to make a decision. Therefore, a voting procedure consists of determining from one method the winner(s) of a vote. Thus, there are several voting procedures that have emerged depending on the specific situations. We can cite: the majority with its variants, the approval or assent, the weighted, elimination, Borda, Condorcet [6,7]. The two most common methods are Borda and Condorcet and the following points describe their basic principles or rules [8,9].

The Borda method

If we have $\boldsymbol{n}$ candidates or alternatives, Borda's method works as follows: Each voter classifies alternatives following his preference order. For each ballot cast, points are awarded to each alternative as follows: The first-place ranking is worth $\boldsymbol{n}$ points, the 2^{nd} place ranking is worth $\boldsymbol{n-1}$ points, and so on. The last place ranking is worth **1** point. The alternatives are then ranked in the preference order by how many points they have accumulated. And the alternative that gets the highest points is the winner. For example: seven (7) voters 1, 2, 3, 4, 5, 6, 7, ordering three (3) candidates, A, B, C, as follows

Voter→		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**Points**
Ranking	1th	C	C	B	B	A	A	A	3
	2nd	B	B	C	C	B	C	C	2
	3th	A	A	A	A	C	B	B	1

P(A)= (3*3) + (4*1) = 13 points
P(B)= (2*3) + (3*2) + (2*1) = 14 points
P(C)= (2*3) + (4*2) + 1 = **15 points**
Hence C is the winner with 15 points

Table 1: Example of Borda

The Condorcet method

Each voter classifies alternatives following his preference order. To prevail, an alternative must undo the others in a "one by one" confrontation. One compares the alternatives by raising the number of votes obtained by duel. The winner is the one that has won its duels against the other alternatives. Let us use the same example and taking ">" as a representation of the preference relation of one candidate to the other one obtains

Duel→	A versus B	A versus C	B versus C
	A>B = 3	A>C = 3	B>C = 3
	B>A = 4	C>A = 4	C>B = 4
Winner→	B	C	C

C won all his duels, therefore he is the winner.

Table 2: Example of Condorcet

Existing voting tools

Several tools are developed, some of them are free and other are not. Among them we can refer to Condorcet Vote [10], Decision Maker [11,12], RoboVote [13], Tricider [14], Whale3 [15], Electionrunner [16], VoxVote [17], Doodle [18], etc. Two of them are largely used and the following section describes them.

RoboVote is a free service that helps users to combine their preferences or opinions into optimal decisions. To do so, RoboVote employs state-of-the-art voting methods developed in artificial intelligence research. This solution offers two types of polls (Objective Opinions and Subjective Preferences), which are tailored to different scenarios; it is up to users to indicate to RoboVote which scenario best fits the problem at hand. Its algorithms are based on the Condorcet's voting method [13].

Whale3 (WHich ALtervative is Elected) It is a web application dedicated to collective decision-making and voting. You can create a poll (open or secret), invite people to participate, and view the results. Whale3 offers several modes of preference expression (ordinal, qualitative, binary, numerical), and relies on voting theory to illuminate the results [15].

VOTING TOOLS IN GRUS

GRUS is equipped with several collaborative tools and can be used in various decision-making situations: asynchronous or not and distributed or not. Depending on the contexts, GRUS allows to define the stages of the decision-making process of a group. Each stage may correspond to the use of a specific tool. For example, a decision-making process of a group can be composed of the following steps: - Brainstorming - Categorization of ideas - Establishment of consensus. In the case of a consensus, it will be easy to conclude. If not, a recourse to voting procedures is necessary.

A voting plugin tool for GRUS will have the following particularities in its use: a) ability to interact, in the processes of a meeting, with the other tools existing on GRUS; b) usable alone as a vote planner like the other solutions explained in the **section of Existing voting tools**. To achieve this, it follows the following logical architecture:

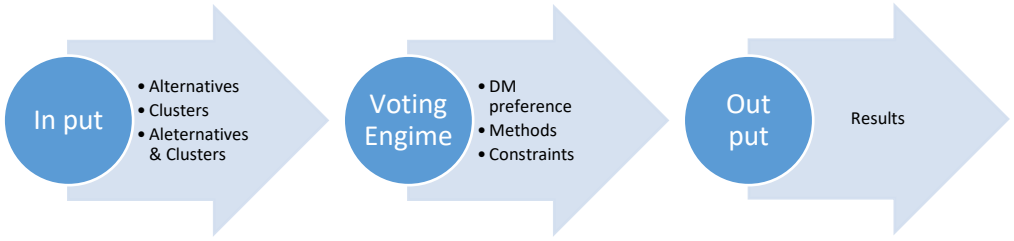


Figure 1: Logical architecture of GRUS's plugin voting tool

Inputs may vary from case to case. Alternatives, clusters of alternatives or a mixture of the two can be used as input for the vote during a meeting. The Voting Engine, receives these inputs, applies one or more methods according to the preferences of the voters for a proclamation of the result which leads to the outputs.

The current status of the implementation

Currently, we have developed the Borda method in the voting tool for GRUS. And we'll use a practical example to assess the current state of implementation.

Suppose that in a management meeting of a company, five (5) leaders must decide on the adoption of a new source of energy for the respect of the environment. The alternatives selected are: wind, solar or hydroelectric among several proposed. So we will have five (5) voters ($V_1, .. V_5$) who must choose between three (3) alternatives ($Alternative_1, Alternative_2, Alternative_3$) corresponding respectively to wind, solar and hydroelectric. The Table3 represents the actual preferences of voters.

Ranking	V_1	V_2	V_3	V_4	V_5
1th	Wind	Solar	Hydroelectric	Solar	Wind
2nd	Hydroelectric	Wind	Solar	Hydroelectric	Solar
3th	Solar	Hydroelectric	Wind	Wind	Hydroelectric

Table 3: Preference of five voters

With the GRUS voting tool, the facilitator must do the following:

a) Create a vote that takes into account the data of the example as shown in Figure 2.

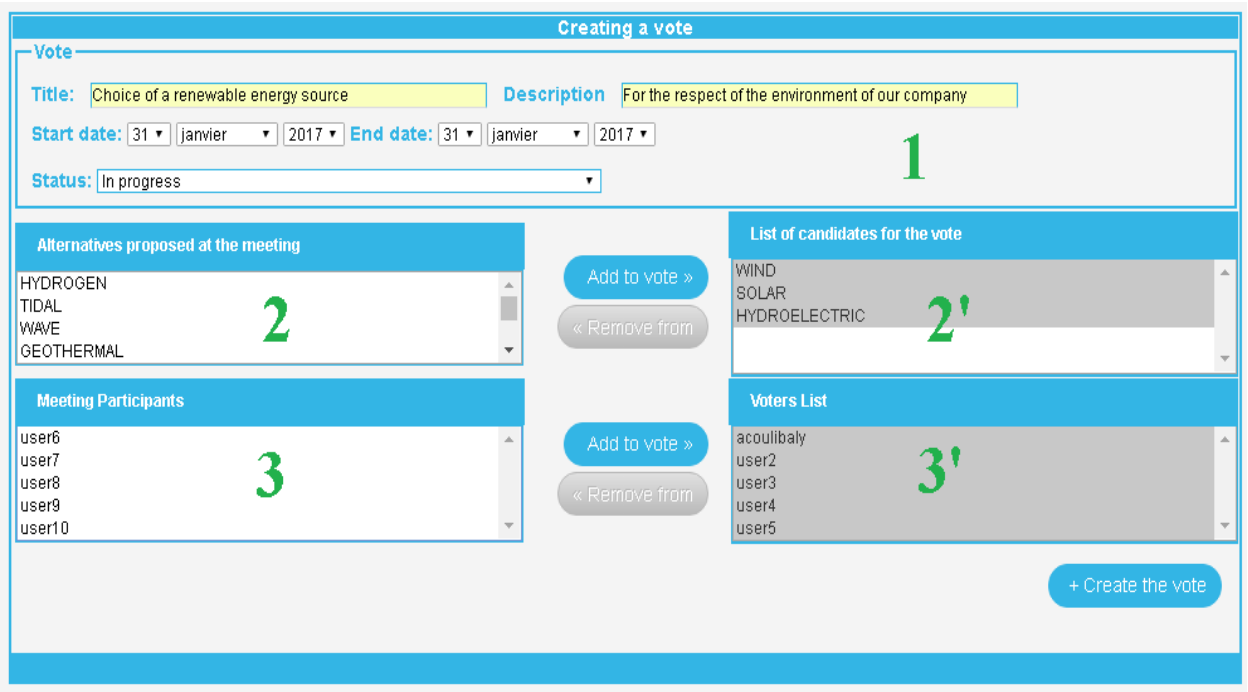


Figure 2: Voting creation Screen

1: descriptions of the vote to know: title, description, start and end dates, status
2: the alternatives proposed in the meeting. 2': the selected alternatives used for the vote
3: The participants of the meeting. 3': the selected participants who will vote.

b) Let the participants vote: Expression of the preferences of each voter as shown in Figure 3.

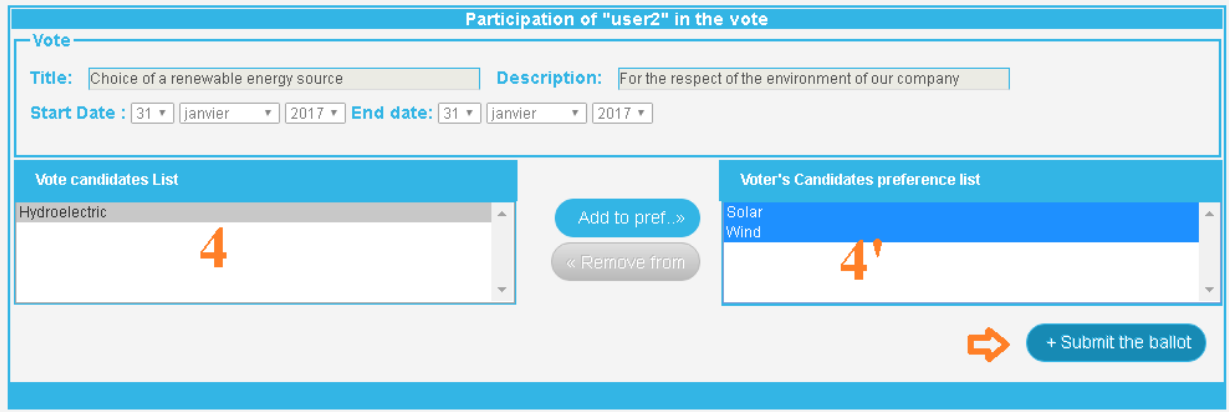


Figure 3: Vote participation screen

4: the list of candidates for the vote (2 ')
4': the list of preference of the voter representing his choices on the alternatives.

c) Proclaim the result of voting (see Figure 4)

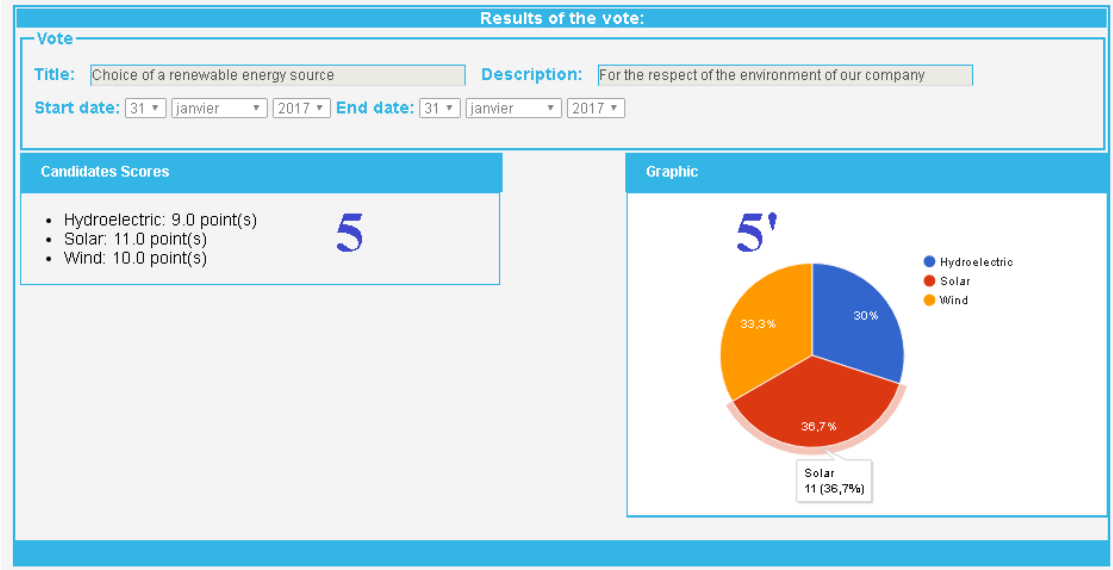


Figure 4: Voting results screen

5: numerical scores of the candidates.
5': graphical representation of the voting result
So the Solar with 36,7% is chosen by the management team.

This example shows that a first implementation has been done. Because it contains features such as management of vote (create, update, delete), participation in a vote and the announcement of the result of applying the method Borda.

CONCLUSIONS AND PERSPECTIVES

There are several voting procedures that are well documented in the literature because they have been the subject of a lot of researches. They are also present in many tools, some of which are free. A voting tool plugin for GRUS, gives it a considerable advantage to be a very complete GDSS.

In its state, this plugin is functional and uses the Borda method, for the calculation of the votes result. This method makes it possible to choose one or more alternatives generally benefiting a high degree of satisfaction of the voters.

As in perspectives of this work, we propose to:

- implement other voting procedures in the GRUS's plugin voting tool;
- conduct a detailed study of the impact of meeting contexts in order to implement the recommendation feature in this voting tool;
- integrate the plugin into the GRUS's toolbox.

REFERENCES

1. A. Abdelkader, « Aide à la Facilitation pour une prise de Décision Collective : Proposition d'un Modèle et d'un Outil » Université de Toulouse, 2010.
2. P. Zaraté, Outils pour la décision coopérative, Toulouse : Hermès, 2013.
3. G. L. Kolfschoten, M. Den Hengst-Bruggeling, G.J. De Vreede, «Issues in the Design of Facilitated Collaboration Processes » Group Decision and Negotiation, pp. 16, 347-361., 2007.
4. P. Zaraté, D.M. Kilgour et K. Hipel, «Private or Common Criteria in a Multi-Criteria Group Decision Support System! An Experiment» Collaboration and Technology: 22nd International Conference, CRIWG, 2016.
5. S. J. Brams, D. M. Kilgour, «Narrowing the field in elections: the next-two rule» Journal of Theorical Politic, pp. 24,4,504-525, 2012.
6. D. Lepelley, H. Smaoui, « Choix collectif et procédures de vote » Eco Flash, pp. 1-20, 2012.
7. D. Bouyson, T. Marchant, P. Perny, « Théorie du choix social et aide multicritère à la décision » pp. 1-40, 2005.
8. A. T. De Almeida, H. Nurmi, «Aiding the choice of a voting procedure» GDSS2014, 2014.
9. S. J. Brams, P. C. Fishburn, Handbook of social choice and welfare: chap4 Voting procedure, Elsevier, 2002, pp. 175-227.
10. J. Boudry, « Condorcet Vote ». Available: https://www.condorcet.vote/. [Access 01/ 2017].
11. «Decision Maker» de Borda Institute, [En ligne]. Available: http://www.decision-maker.org/content/voting-systems. [Access 01/2017].
12. De Borda Institute, «Voting system from de Borda counting». Available: http://www.deborda.org/. [Access 11/2016].
13. Carnegie Mellon University, Harvard University, «RoboVote». Available: http://www.robovote.org/. [Access 01/ 2017].
14. N. Graf von Kanitz, S. Eyl, « Tricider ». Available: https://www.tricider.com/. [Access 01/ 2017].
15. « Whale3 ». Available: http://strokes.imag.fr/whale3/. [Access 01/ 2017].
16. « Electionrunner ». Available: https://electionrunner.com/. [Access 01/ 2017].
17. V. Van Witteloostuyn, «VoxVote » Available: http://www.voxvote.com/. [Access 01/ 2017].
18. Doodle, « Doodle - Le vote en ligne ». Available: https://doodle.com/fr/vote-en-ligne. [Access 01/2017].

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

Negotiation style and cognitive effort in a Negotiation Support System: a biometric analysis

Jadielson Alves de Moura, Suzana de França Dantas Daher and Ana Paula Cabral Seixas Costa
Universidade Federal de Pernambuco, CDSID - Center for Decision Systems and Information Development;
Av. da Arquitetura, S/N, Cidade Universitária, Recife, 50740-550, Brazil
jadielson.moura@gmail.com, suzanadaher@cdsid.org.br, apcabral@hotmail.com

ABSTRACT

This study aims to investigate the relationship between the cognitive effort and negotiation style during negotiation processes by conducting an experiment using a Negotiation Support System (NSS). An eye tracker device is used to gather psychophysiological data from participants. The pupil size is widely known as a reliable metric for measuring the cognitive effort, and so could provide information about stress level, while people interacts with information system. Participants performed a simulated negotiation task in an NSS environment. All participants were grouped in accordance with their negotiation style, operationalized as levels of assertiveness and cooperativeness. Negotiators demonstrated significant differences in pupillary variation and consequently in cognitive effort during the trial. These results suggest that the negotiation style of the negotiator can indicate the possible recognizing pattern adopted and their mental workload spent in negotiation situations supported by the NSS.

Keywords: Cognitive effort, Negotiation style, Eye-tracking, Negotiation Support System.

INTRODUCTION

Negotiation can be understood as a social interaction process in which people mutually allocate resources by negotiating with each other during which they seek to resolve their opposing interests, and to make joint decisions [1][2]. This interactive process is widely used in studies of different natures, and this includes both in situations involving face-to-face and/or computer-mediated negotiations. In particular, computer-mediated negotiation is characterized as an interactive communication and decision-making process between at least two individuals using an electronic system [3]. Although simple tools such as email could be used in an asynchronous negotiation, we analyze a negotiation which is supported by a specific online negotiation platform known as the Negotiation Support System (NSS) [4].

Although the NSS provides technological resources for supporting the negotiation process, the process for using it often becomes stressful due to the time that needs to be spent to reach an agreement and the large number of interactions between the negotiators. These negative factors can be increased when the negotiation process is driven only in online interactions, without physical or face-to-face contact. The amount of information exchanged in each interaction, as well as social contagion, have been factors that have influenced the time taken to complete negotiations and the way in which negotiators interact [5]. Therefore, these matters may positively or negatively affect the negotiators' levels of stress.

The mutual influences and the level of stress result from negotiators exchanging information with each other in conjunction with social contagion may guide the negotiator's behavior during the negotiation process [6]. This behavior - which can be described in terms of the negotiation style or of conflict management [7] - often changes while negotiators interact with other parties in the negotiation. Such changes in behavior are related to the weak relationship between what is stated by participants and their negotiation style in practice [8]. In addition, the influences arising from the emotional transitions commonly found in social interactions also take place during negotiations [9].

From the context started by previous research studies [8][9], this paper sets out to investigate the role of negotiation style in the behavior of negotiators, while interacting with the informational tool, the NSS, in terms of element recognition and cognitive workload expended in order to understand the information depicted. To do so, desktop eye tracker equipment with a 120 Hz sampling rate is used to capture measures of negotiators' stress levels by recording as a metric the variation in size of the negotiators' pupils. To capture the initial negotiation style, the Thomas-Kilmann Conflict Mode Instrument (TKI) is used [7].

The TKI model was applied in several contexts of problem resolution to help individuals understand their interpersonal relationships across culture, gender, organizational role, and so on [10][11][12]. The negotiation area also applied the TKI to conflict resolution issues [8][10]. In addition, TKI approaches were proposed for negotiation supported by NSS, such as Jain and Solomon's (2000)[13] study.

METHOD

The experiments for testing the assumptions of this paper have been conducted in a specific module of NegPlace System which is a web-based NSS developed by Moura & Costa [14]. Initially, participants answer a questionnaire to capture their individual negotiation style. Forty students from Universidade Federal de Pernambuco - 14 undergraduate and 26 post-graduate students (mean age = 26.45, SD = 4.42) - took part in the study on a voluntary basis. The number of males (N = 17) and females (N = 23) are not equal,

but a previous study has shown that gender does not influence a participant's pupillary response [15]. During the experiment was not imposed any time restriction, since it is not wanted to press participants.

The dimensions of each participant´s pupils are gathered during all experiments and analyzed together with other psychophysiological metrics and data on the negotiators' behaviors. A statistical analysis is performed to find the relationship between the cognitive workload level and negotiation style during negotiation processes.

RESEARCH MODEL AND RESULTS

Although individuals are divided into negotiation style groups per their levels of assertiveness and cooperativeness, the same negotiation task and informational tool was available to each negotiator. Thus, we used eye-tracking measures as well as pupillometry to provide scanpaths of the Areas of Interest (AOI) and cognitive workload according to negotiation style.

All participants were assigned the role of the seller, in which they needed to sell their car and have announced it in an online car sales website. Thus, they had received a counteroffer proposed by anonymous buyer that is waiting for a reply, accepting the offer or proposing a counteroffer. An overview with some information about personality traits and negotiation styles of the anonymous buyer were available at the information tool in the NSS.

The data collected were subjected to two types of analysis, one for Area of Interest (AOI) and another for eye movement. In order to carry out AOI analysis we defined five AOIs at the Informational tool screen. The metrics *time to first fixation, visit count,* and *total visit duration for each AOI* were calculated for each participant, and then calculated the average for each group. At the last of AOIs analysis, each AOI was appointed to one group either *graphical* or *textual* group, in accordance to the type of element in it. Pupillary responses were analyzed using Windows SPSS 22.

A descriptive analysis was used to find out which AOIs were more visited by the eyes and for how long. In addition, the *time to first fixation* of AOIs was also analyzed to clarify the average time spent to fix attention on a specific area. All measures were calculated for all participants as well as for each group: group 1 (partially assertive and cooperative), group 2 (more assertive and less cooperative), and group 3 (less assertive and more cooperative). The AOIs description data is summarized in Table 1.

Table 1. AOIs Descriptive Analysis

AOIs		**N**	**Mean**	**Std. Deviation**	**Std. Error**
Total Visit Duration in Graphical Information (second)	Group 1	15	7.13	6.36	1.64
	Group 2	16	6.41	5.15	1.28
	Group 3	9	6.69	8.73	2.91
	Total	40	6.74	6.36	1.00
Total Visit Duration in Textual Information (second)	Group 1	15	18.18	16.30	4.20
	Group 2	16	21.49	14.80	3.70
	Group 3	9	22.48	18.83	6.27
	Total	40	20.47	15.99	2.52
Visit Count in	Group 1	15	11.60	6.93	1.79

Graphical Information	Group 2	16	13.00	9.52	2.38
	Group 3	9	11.77	11.76	3.92
	Total	40	12.20	8.99	1.42
Visit Count in Textual Information	Group 1	15	19.20	10.79	2.78
	Group 2	16	21.12	13.84	3.46
	Group 3	9	15.44	11.41	3.80
	Total	40	19.12	12.12	1.91

An ANOVA One-way was performed for both left and right pupil size with groups' pupil size means as a factor analysis. The groups' means were the results of the temporal dynamics of the group's pupillary response during the first ten seconds (10000 ms) for each participant in the experiment, divided into 40 bins of equal size (250 ms). The Table 2 displays the variance analysis that revealed the groups differed significantly on the pupil size means for both left [$F(2, 1154) = 35.79$, $p < .0001$] and right [$F(2, 1154) = 59.61$, $p < 0.0001$]. These results indicate that cognitive load evoked during the informational tool interactions differs according to cooperative and assertive levels of the clustered groups.

Table 2. ANOVA One-way for left and right pupil size between groups

		Sum of Squares	**df**	**Mean Square**	**F**	**Sig.**
Pupil Left	Between Groups	18.098	2	9.049	35.793	0.000
	Within Groups	291.238	1152	0.253		
	Total	309.336	1154			
Pupil Right	Between Groups	28.813	2	14.406	59.618	0.000
	Within Groups	278.371	1152	0.242		
	Total	307.184	1154			

A post-hoc analysis revealed a larger pupillary response to Group 2 compared to other groups especially from fourth seconds onwards (4000 - 1000 ms) with a peak between eight and nine seconds. The Group 3 has a moderate pupillary response compared to Group 1 from the onset until seventh seconds. After seven seconds onwards, the pupillary response to Group 3 follow closely to Group 1, sometimes the size slightly larger. The Group 1 has the lowest pupil size mean, which remains close to 3.1 mm of pupil diameter for both left and right pupil during the course of the trial. However, in two instances for both pupils, the pupil size mean to Group 1 is greater than or equal to the pupil size mean of the Group 3.

To sum up, results have shown that cognitive workload on interaction with informational tool differs from group to group, depending on the level of assertiveness and cooperation. Thus, the group of participants that are more assertive and less cooperative presents a high cognitive load on interacting with informational tool content, if compared to other groups. In its turn, the participants of the group that is partially assertive and cooperative tends to vary somewhat their cognitive load along the trial. While the participants of the group that are less assertive and more cooperative are prone to decrease their cognitive load over time.

CONCLUSIONS

The present study investigated the relationship between the cognitive effort and negotiation style during negotiation processes by conducting an experiment using a Negotiation Support System, so called NegPlace. The negotiators were clustered according their levels of assertiveness and cooperativeness from the negotiation style into three groups: more assertive and less cooperative; partially assertive and cooperative; and less assertive and more cooperative. Consistent with our expectations, negotiator groups differed in relation to recognizing a sequence of AOIs and average pupil size. Differences in eye-movement when recognizing AOIs could be due to the high assertive level, which is related to attempting to satisfy one's own concern, avoiding qualitative issues, and focusing on quantitative issues that can yield substantial value [10]. This might have motivated their fastest interest in the negotiation information of the opponent than would have other groups, while the negotiator with a high cooperative level attempts to satisfy the concerns of the other negotiator and engaging in problem-solving behavior [8].

In the observed pupil size results, we noted that the cognitive effort of the negotiators differs significantly between groups. The pupil size patterns demonstrate that the more assertive and less cooperative group required more cognitive processing than other groups. This result reveals the efforts of the group in recognizing and interpreting information in the informational tool. These findings could support the development of better NSSs interfaces. These results highlight the psychological and behavioral differences in negotiating styles (assertive and cooperative) through physical measures, demonstrating the application and relevance of psychological models such as Thomas-Kilmann Conflict Mode Instrument (TKI) in negotiating issues, using the NSS.

One of the limitations of this study is not established a fixed time and scanpaths in executing the task, which provide a high variance at completion time of trial and several scanpath patterns. Other limitations are notable. First, the sample was homogeneous in age and course, which limits generalizability. Second, the negotiation task has only one counteroffer, limited to a single interaction. Third, no real rewards were offered to participants. Further research is needed to investigate negotiation style pattern while interacting with an information tool in other negotiation problems, such as collaborative negotiations.

REFERENCES

[1] D. G. Pruitt, Negotiation Behavior. New York: Academic Press, 1981.

[2] L. Thompson and R. Hastie, "Social perception in negotiation," Organ. Behav. Hum. Decis. Process., vol. 47, no. 1, pp. 98–123, Oct. 1990.

[3] M. Schoop, "The Worlds of Negotiation," Proc. 9th Int. Work. Conf. Lang. Perspect. Commun. Model., vol. 2004, no. Lap 2004, pp. 179–196, 2004.

[4] T. Jelassi and A. Foroughi, "Negotiation support systems: an overview of design issues and existing software," Decis. Support Syst., vol. 5, no. 2, pp. 167–181, Jun. 1989.

[5] L. L. Thompson, "Negotiating via Information Technology," Mind Hear. Negot., vol. 58, no. 1, pp. 319–337, 2008.

[6] M. Gomes, T. Oliveira, D. Carneiro, P. Novais, and J. Neves, "Studying the Effects of Stress on Negotiation Behavior," Cybern. Syst., vol. 45, no. 3, pp. 279–291, 2014.

[7] K. W. Thomas and R. H. Kilmann, Thomas-Kilmann Conflict Mode Instrument. New

York, NY: Xicom, Tuxedo, 1974.
[8] O. Miller, “The Negotiation Style: A Comparative Study between the Stated and in-Practice Negotiation Style,” Procedia - Soc. Behav. Sci., vol. 124, pp. 200–209, Mar. 2014.
[9] A. Filipowicz, S. Barsade, and S. Melwani, “Understanding emotional transitions: The interpersonal consequences of changing emotions in negotiations.,” J. Pers. Soc. Psychol., vol. 101, no. 3, pp. 541–556, 2011.
[10] G. R. Shell, “Bargaining styles and negotiation: The Thomas-Kilmann conflict mode instrument in negotiation training,” Negot. J., vol. 17, no. 2, pp. 155–174, 2001.
[11] S. D. Brahnam, T. M. Margavio, M. a. Hignite, T. B. Barrier, and J. M. Chin, “A gender-based categorization for conflict resolution,” J. Manag. Dev., vol. 24, no. 3, pp. 197–208, 2005.
[12] O. Gbadamosi, A. Ghanbari Baghestan, and K. Al-Mabrouk, “Gender, age and nationality: assessing their impact on conflict resolution styles,” J. Manag. Dev., vol. 33, no. 3, pp. 245–257, 2014.
[13] B. a. Jain and J. S. Solomon, “The effect of task complexity and conflict handling styles on computer-supported negotiations,” Inf. Manag., vol. 37, pp. 161–168, 2000.
[14] Moura, J., Costa, A.P.C.S. "NegPlace a Web Negotiation Support System that incorporates negotiators' styles". In: 20th Americas Conference on Information Systems, 2014, Savannah. AMCIS 2014 PROCEEDINGS, 2014, v. 1, 2014
[15] J. Beatty, “Task-Evoked Pupillary Responses, Processing Load, and the Structure of Processing Resources,” Psychol. Bull., vol. 91, no. 2, pp. 276–292, 1982.

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

ICDSST 2017 on Formulating Dynamic Data Replication Strategies in Cloud Systems: A Survey, Issues, and Future Research Directions

Sean Eom
Southeast Missouri State University
Cape Girardeau, MO 63701, USA
sbeom@semo.edu
http://cstl-hcb.semo.edu/eom/

ABSTRACT

Data replication is making multiple copies of part of data in an organization and placing them at multiple sites. Proper formulation of data replication strategies is a subset of a broad topic of management of geographically distributed, heterogeneous computing and storage network resources. Replica placement and replica factor have been on-going challenges when formulating a replica strategy in the cloud environment. In this paper, first, we present a survey of data replication strategy classification in cloud systems. Second, the survey we have conducted reveals that there are several open issues with the current data replication strategies. We will discuss the open issues. Finally, we present a new research framework based on the two perspectives: multiple criteria decision making and information resource management perspectives.

Keywords: Data Replication, Distributed Computing, Heterogeneous Computing, Multiple Criteria Decision Making, Cloud Computing, Information Resource Management.

INTRODUCTION

Data replication is a widely used technique in cloud systems to improve data availability and to minimize the bandwidth consumption and data access latency. It is frequently used in parallel and distributed systems, large scale systems, including P2P [1] and data Grid systems [2]. Many replication strategies aim to answer the following questions: what data should be replicated; when to replicate; and where the new replicas should be placed.

Most of works in the literature have classified replication strategies in classical environments, according to several criteria: (i) static vs. dynamic classification, (ii) centralized vs decentralized replication management, (iii) server vs. client replication, (iv) objective function and (v) system architecture. (For a detailed review and discussion of the various classification, readers are referred to [3].

Dynamic replication strategies automatically create and remove replicas based on the changes in user access pattern, storage capacity and bandwidth. Bai et. al. [4] proposed a dynamic Response Time-based Replica Management (RTRM) strategy. This strategy consists of three levels: (i) replica creation, (ii) replica selection and (iii) replica placement. Simulation results showed that RTRM strategy improves network utilization and service response time. Another strategy, Cost-effective Dynamic Replication Management (CDRM) is proposed to capture the relationship between replica number and availability. It aims to provide cost-effective data availability while improving both the load balancing and system performance of the cloud (Wei et al 2010). Readers are referred to [5] for a comprehensive review of static vs, dynamic data replication strategies. Due to the dynamic nature of cloud systems, dynamic replication strategies seem to be more suitable in cloud systems.

The first part of this paper presents a survey of data replication strategies in cloud systems. Second, the survey we have conducted reveals that there are several open issues with the current data replication strategies. We discuss the open issues. Finally, we present a new research framework and future research directions, based on the two perspectives: multiple criteria decision making and information resource management perspectives.

A SURVEY OF DATA REPLICATION STRATEGIES

Recently, Tabet and others [3] discuss existing classifications and how dynamic replication can be classified with respect to many criteria. They point out the advantages and weakness of each of the existing strategies present a new taxonomy of existing data replication strategies incorporating multiple dimensions of factors to consider. The multiple dimensions of factors they used include:

(i) Static vs. Dynamic [4]
(ii) Reactive vs. Proactive workload balancing solutions [6]
(iii) Provider vs. Customer centric [7, 8]
(iv) Replica factor determination [9] and
(v) The objective function including data locality, network bandwidth locality, and replication costs [10, 11].

The survey of Tabet et al. [3] is most recent survey of data replication strategy encompassing a wide range of multiple dimensions of factors. Nevertheless, this survey does not satisfactorily provide solutions to open issues. The next section discusses several open issues as a basis of our comprehensive future research directions.

OPEN ISSUES IN DYNAMIC REPLICATION STRATEGIES

Milami & Navimipour [5] presented several open issues in data replication in the cloud environment. Therefore, they believe that designing the data replication strategies that considers the important parameters/multiple criteria of data replication problems in a cloud environment is a challenging task.

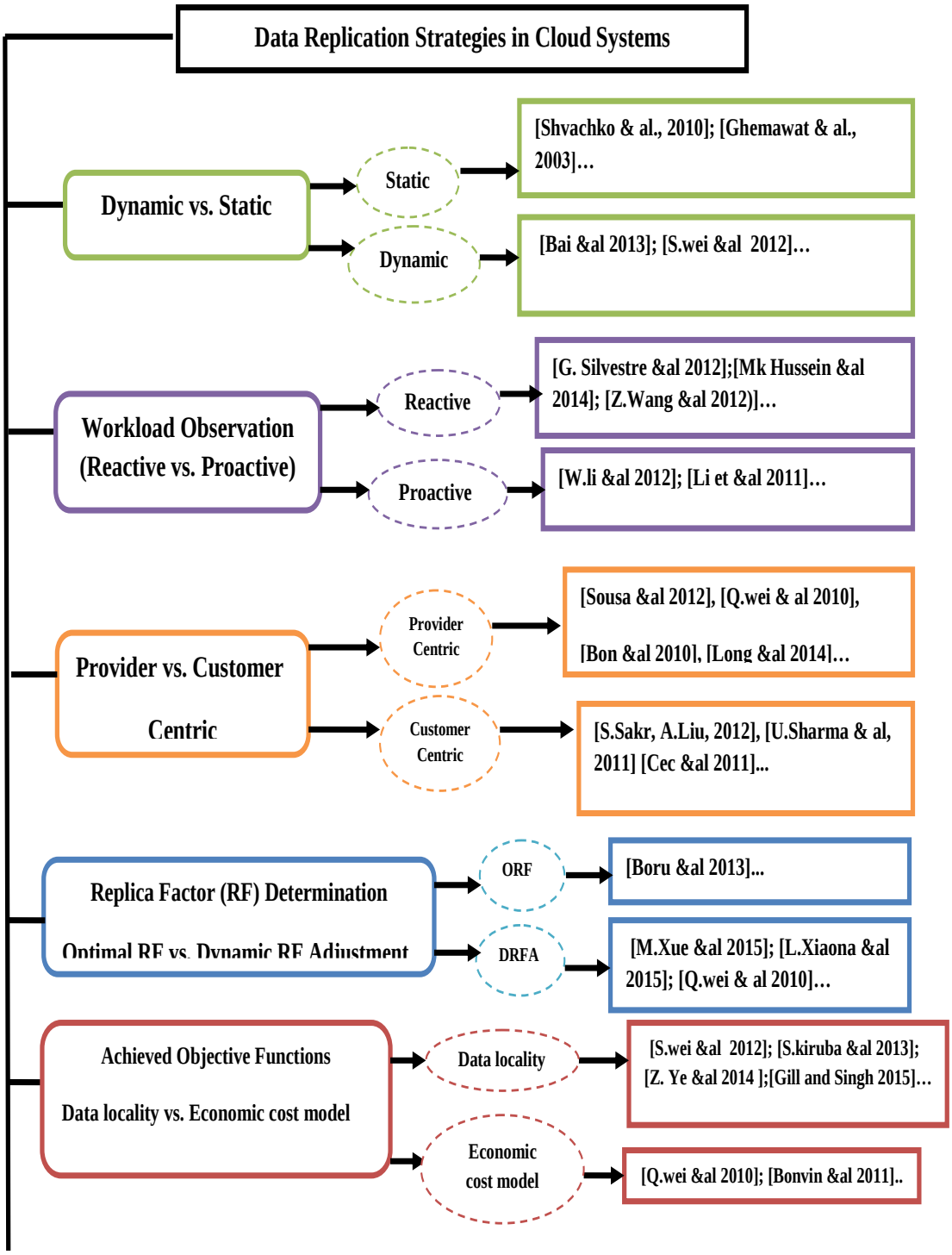


Figure 1: Classification of data replication strategies in cloud systems
Source: [3]

1. Most of studies they reviewed used a hierarchical structure. Therefore they suggest that it may be very interesting to modify the hierarchical structure to reflect the real cloud environment.
2. No single strategies can address all issues involved in the data replication including reliability, scalability, fault tolerance, user waiting time, data access time, conserving the network bandwidth, and load balance.
3. It may be desirable to improve selection process to allow users to elicit their preferences as to the objectives of replication strategies including data security, data reliability, etc.[12]

Recently, Tos et al. [12] provided a new taxonomy of dynamic replication strategies that takes the impact of data grid architecture on dynamic replication performance into consideration. The issue no. 1 raised by Milami et al. [5] is effectively tackled by Tos et al. They present various replication strategies for different data grid architectures: hierarchical, peer-to-peer, hybrid, general graph. They concluded based on a simulation study that more relaxed architectures such as general graphs produced better response times. They also discussed open issues in dynamic replication strategies.

4. Many replication strategies aim to accomplish multiple conflicting objectives, including improving performance/availability vs. increasing cost burdens to providers as well as the consumers.
5. Studying the combined effectiveness of the optimal number of replicas and strategically placing them is open issues.

FUTURE RESEARCH DIRECTIONS: RESOURCE MANAGEMENT AND MCDM PERSPECTIVES

Of these five open issues in the literature, the first issue is not valid any more since it was effectively tackled by Tos et al [12]. Careful analysis the remaining four issues can be boiled down to the creation of a strategy that effectively deals with multiple criteria including reliability, scalability, fault tolerance, user waiting time, etc. and allowing the user to take part in the formulation process to express their preferences with priorities. To do so, our suggestion for future research direction is taking a broader view of the resource management and multiple criteria decision making perspectives.

The Resource Management Perspective: Data replication is a subset of resource management in the cloud. The resource management in the cloud is described as the process of allocating computing, storage, networking and (indirectly) energy resources to a set of applications, in a manner that seeks to jointly meet the performance objectives of the applications, the IT infrastructure providers and the users of the cloud resources, within the constraints of Service Level Agreements (SLAs) with the Cloud Users. These objectives of different entities are clearly conflicting. [13, pp. 568-569]. Accordingly, data replication strategy must be formulated with several objectives of data resource management. Scope of cloud computing resource management [13] include the following:

Actors: they include cloud providers whose roles are managing a set of data center hardware and system software to meets SLAs agreed with Cloud Users/End Users and/or achieves other management goals, Cloud users who are responsible for meeting SLAs agreed with its

customers (i.e., End Users) with the objectives of minimizing its costs and maximizing its profits.

Management Objectives of the Cloud Provider aiming to meet service level agreements (SLAs) with Cloud Users.

Management Objectives of The Cloud User aiming to exploit the elasticity property of cloud environments.

Resource Management Functions: Resource management functions with regard to data replication cover the following:

- Resource Demand Profiling (effective resource management must continuously monitor end-user demand of network resources and achieve an appropriate balance between the current demand (reactivity) and predicted demand (proactivity).

- Work Load Management Goal to ensure that the benefit of a given strategy is higher than the cost of replication [14].

The Multiple Criteria Decision Making Perspective:

All open issues introduced above are clearly involved with multiple criteria decision making (MCDM) which inherently necessitates a simultaneous comparison of the large number of criteria and alternatives that, in turn, demands the more complex array of information. These features include (1) the multiple objective structure designed to handle both quantitative and qualitative information crucial for ill-structured problems. In addition, the emphasis on the decision maker's judgment or bounded rationality which better reflects the decision maker's actual cognitive behaviors. Therefore, it is clear that taking the MCDM perspectives allow users to elicit their preferences as to the objectives of replication strategies including data security, data reliability, etc.

CONCLUSIONS

Information systems researchers have continuously struggled to deal with multiple criteria decision problems over the past several decades [15]. The taxonomy we presented can be a useful guideline for IT managers to select the data replication strategy for their organization. As previously stated [5], we do not believe that there is an optimal data replication strategy that addresses all of the issues involved in the data replication. Futhermore, we believe that it is highly unlikely to find such an ideal replication strategy that optimizes a wide range of goals of actors (providers and customers) and organizational policies. Rather, the replication strategy for any organization will be a decision that satisfice several goals of the entities involved. It may be a very promising direction for future researchers to approach the issues in the replication strategies from two different perspectives: multiple criteria decision making and resource management in the cloud.

REFERENCES

1. F. Xhafa, V. Kolici, A.-D. Potlog, E. Spaho, L. Barolli, and M. Takizawa. "Data Replication in P2P Collaborative Systems". in *The 7th International Conference on*

P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC). Victoria, BC: IEEE. 2012.
2. Y. Mansouri, S.T. Azad, and A. Chamkori. "Minimizing Cost of K -Replica in Hierarchical Data Grid Environment". in *2014 IEEE 28th International Conference on Advanced Information Networking and Applications*. IEEE. 2014.
3. K. Tabet, R. Mokadem, M.R. Laouar, and S.B. Eom, "Data Replication in Cloud Systems: A Survey". International Journal of Information Systems and Social Change (IJISSC), 8(3). 2017.
4. X. Bai, H. Jin, X. Liao, X. Shi, and Z. Shao, "RTRM: A Response Time-Based Replica Management Strategy for Cloud Storage System", in *Grid and Pervasive Computing: 8th International Conference, GPC (Grid and Pervasive Computing) 2013 and Colocated Workshops, Seoul, Korea, May 9-11, 2013. Proceedings* J.J.J.H. Park, et al., Editors. Springer Berlin Heidelberg. p. 124–133. 2013.
5. B.A. Milani and N.J. Navimipour, "A Comprehensive Review of the Data Replication Techniques in the Cloud Environments: Major Trends and Future Directions". Journal of Network and Computer Applications, 64: p. 229–238. 2016.
6. M.-K. Hussein and M.-H. Mousa, "A Light-Weight Data Replication for Cloud Data Centers Environment". International Journal of Engineering and Innovative Technology (IJEIT), 1(6): p. 169-175. 2014.
7. F.R.C. Sousa and J.C. Machado. "Towards Elastic Multi-Tenant Database Replication with Quality of Service". in *Proceedings of 2012 IEEE/ACM Fifth International Conference on Utility and Cloud Computing* IEEE. 2012.
8. S. Sakr and A. Liu. "Sla-Based and Consumer-Centric Dynamic Provisioning for Cloud Databases". in *2012 IEEE 5th International Conference on Cloud Computing (CLOUD)*. IEEE. 2012.
9. M. Xue, J. Shen, and X. Guo. "Replica Placement in Cloud Storage Based on Minimal Blocking Probability". in *CENet2015* Shanghai, China. 2015.
10. S. Kirubakaran, S. Valarmathy, and C. Kamalanathan, "Data Replication Using Modified D2RS in Cloud Computing for Performance Improvement". Journal of Theoretical and Applied Information Technology 58(2): p. 460-470. 2013.
11. U. Tos, R. Mokadem, A. Hameurlain, T. Ayav, and S. Bora, "A Performance and Profit Oriented Data Replication Strategy for Cloud Systems", in *The IEEE International Conference on Cloud and Bigdata Computing* Toulouse. 2016.
12. U. Tos, R. Mokadem, A. Hameurlain, T. Ayav, and S. Bora, "Dynamic Replication Strategies in Data Grid Systems: A Survey". The Journal of Supercomputing, 71(11): p. 4116–4140. 2015.
13. B. Jennings and R. Stadler, "Resource Management in Clouds: Survey and Research Challenges". Journal of Network and Systems Management, 23(3): p. 567–619. 2015.
14. M.v. Steen and G. Pierre, "Replicating for Performance: Case Studies.", in *Replication, Theory and Practice*, B. Charron-Bost, F. Pedone, and A. Schiper, Editors. Springer: Berlin. p. 73–89. 2010.
15. S.B. Eom, *Longitudinal Author Cocitation Mapping: The Changing Structure of Decision Support Systems Research (1969-2012)*. Foundations and Trends® in Information Systems. Vol. 1. Hanover, MA: Now Publishers (May 25, 2016). 277–384. 2016.

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

ICDSST 2017 on Data, Information and Knowledge Visualisation in Decision Support Systems

Multicriteria aggregating operator adaptation based on decision making context

FOMBA Soumana[1,2], ZARATE Pascale[2], CAMILLERI Guy[3]

1 : Université des Sciences, des Techniques et des Technologies de Bamako (USTTB) - Hamdallaye ACI 2000 - Rue 405 - Porte 359 -BP E423 - Mali
2 : IRIT – Toulouse University – 2 rue du Doyen Gabriel Marty 31042 Toulouse Cedex 9 – France
3 : IRIT – Toulouse University – 118 route de Narbonne 31062 Toulouse Cedex 9 – France

Soumana.Fomba@irit.fr, Pascale.Zarate@irit.fr, camiller@irit.fr
web-page: http://www.irit.fr/

ABSTRACT

Recommender systems aim to support decision-makers by providing decision advice. We review briefly tools of Multi-Criteria Decision Analysis (MCDA), including aggregation operators that could be the basis for a recommender system. Then we develop a generic multi-criteria recommender system, to support decisions by aggregating measures of performance contained in a performance matrix. To determine a total order of alternatives, the system uses different multicriteria aggregation operators depending on the context of use of the system. Thus, recommendations are calculated using partial preferences provided by the decision maker and updated by the system. An integrated web platform is under development.

Keywords: Recommender System, Choquet Integral, MCDA, Sugeno Integral, Aggregation operator

INTRODUCTION

Recommender systems are designed to help decision-makers make the best possible decisions from a wide range of choices. This process of finding the best solution passes through an inescapable step that is the aggregation of the performances of each alternative according to the preferences of the decision maker. For this purpose, there are several multicriteria aggregation operators of preferences. The choice of the aggregation operator in a decision-making problem is therefore crucial.

We propose in this paper a generic recommender system where the choice of the aggregation operator is implicit and transparent for the user. Among these operators are the weighted sum, the ordered weighted sum, the Choquet integral, the Sugeno integral [1], and so on. These operators are classified into two main categories, namely quantitative and qualitative, when they are respectively a decision problem where assessments are quantitative or qualitative [2]. The choice of an operator in a category uses a collaborative recommendation model and a similarity model between decision problems.

This article is structured as follows: we make a brief presentation of some aggregation operators in the first section. In the second section, we present our generic multicriteria recommender system. The third section deals with the results obtained by this system. And finally, we open the discussion and possible prospects to these works.

NOTATION AND FORMALIZATION OF THE PROBLEM

We have the following data:

- Set of alternatives A={a,b,c..} with |A|=m. The alternatives represent the different solutions or choices available to a decision-maker faced with a problem of decision-making.
- Set of criterions $\mathbb{N}=\{1,2,3,..\}$ with $|\mathbb{N}|$ =n. Criteria can be considered as attributes or characteristics of alternatives. Indeed each alternative is evaluated according to each criterion describing it.
- Numerical values taken by the alternatives for each criterion : $\forall j \in \mathbb{N}$, $\forall a \in A$, $a_j \in \mathbb{R}$
- Set of the profiles of the alternatives which is a set of vectors such that $\forall a \in A$ we associate the vector $a=(a_1,a_2, \ldots, a_n) \in \mathbb{R}^n$

Tableau 1: Representation of the performance matrix

Alternative	**Attr. 1**	**Attr.2**	**...**	**Attr.n**	**Aggregated values**
a	a_1	a_2	…	a_n	A(a)
b	b_1	b_2	…	b_n	A(b)
…	…	…	…	…	…

- Let $\succeq$ be a relation on *X* representing the decision-maker's preference. ($\succeq$ is usually pronounced "at least as good as".) As a binary relation, $\succeq$is usually assumed reflexive. For alternatives *a* and *b*, $a \succeq b$ to mean that *a* is preferred to *b*.

AGGREGATION OPERATORS

In decision support, an aggregation operator is usually used to determine an overall score for an alternative from its local performance on the criteria and the user's preferences over criteria, in order to compare it to other alternatives. With the overall score, a ranking can be established that will guide the decision-maker's decision. In this section, we examine several aggregation operators used in the recommender system. These operators belong to two distinct categories, namely quantitative and qualitative.

The Choquet Integral

Aggregation operators such as weighted sum and OWA are unable to model interactions because they depend on weight vectors. What is needed is a non-additive function that defines a weight, not only for each criterion, but also for each subset of criteria. These non-additive functions can thus model both the importance of criteria and the positive and negative synergies between them. A suitable aggregation operator can be based on the Choquet integral [3, 4] that uses non-additive functions that Sugeno proposed be called fuzzy measures [5]. The Choquet integral is used for quantitative evaluations.

The Choquet integral is defined as follows: Let μ be a fuzzy measure on N. The Choquet integral of $x \in R^n$ with respect to μ is defined by:

$$C\mu(x) := \sum_{i=1}^{n} x_{(i)}[\mu(A_{(i)}) - \mu(A_{(i+1)})] \quad (1)$$

where (.) denotes the permutation of the components of $x = (x_1, \ldots, x_n)$ such that $x_{(1)} \leq \ldots \leq x_{(n)}$. As well, $A_{(i)}=\{(i), \ldots, (n)\}$ and $A_{(n+1)} = \emptyset$.

The Choquet integral gives the possibility to calculate the index of interaction between the criteria and the global importance of each criterion, called the Shapley value. For more information see [6]

The Sugeno integral

Unlike the Choquet integral which uses quantitative evaluations, the Sugeno integral is used for qualitative evaluations. The integral of Sugeno has been introduced in [5]. We consider here the integral of Sugeno in its discrete version, applied to the aggregation of preference. We also consider a totally ordered set L, not necessarily a numerical one, which is called an evaluation scale, and whose minimum and maximum elements are denoted by 0 and 1 respectively. Sugeno integral is defined in relation to a capacity on the set N which is a function:

$\mu : 2^n \rightarrow L$ such as : $\mu(\emptyset) = 0$ and $\mu(N) = 1$, for $I \subseteq J \subseteq N : \mu(I) \leq \mu(J)$.

For any set of criteria $I \subseteq N$, the value of $\mu(I)$ can be interpreted as the degree of importance associated with I.Let a capacitance $\mu: 2^n \rightarrow L$. The integral of Sugeno defined with respect to μ, denoted by S_μ, is expressed in the form

$$S_\mu(y_1, y_2, \ldots, y_n) = \bigvee_{i=1}^{n} \left(y_{(i)} \bigwedge \mu(\{(i), \ldots, (n)\}) \right) \quad (2)$$

where (.) denotes the permutation of the components of $y = (y_1, \ldots, y_n)$ such that $y_{(1)} \leq \ldots \leq y_{(n)}$. By this formula it is seen in particular that μ determines S_μ entirely and uniquely. The problem of learning an integral of Sugeno S_μ can therefore be reduced to that of learning the corresponding capacitance μ.

DESCRIPTION OF THE RECOMMENDER SYSTEM

The proposed recommender system can both treat quantitative evaluations using aggregation operators such as the weighted sum, the ordered weighted sum [7], the Choquet integral, and so on, but also qualitative evaluations using the integral of Sugeno. When creating a decision problem, the user is asked to specify the type of problem choosing whether it is a problem where the evaluations are quantitative or qualitative. The number of criteria involved and the description of each criterion are also asked. Then comes the enumeration of the different alternatives and their score for each criterion. Finally, the user is asked to define a partial order on a subset of alternatives. This partial order is used to define the user's preferences between pairs of alternatives. The whole of this information on the problem of decision support is thus stored in a database while disregarding the aggregation operator to be used for the search for a possible better solution.

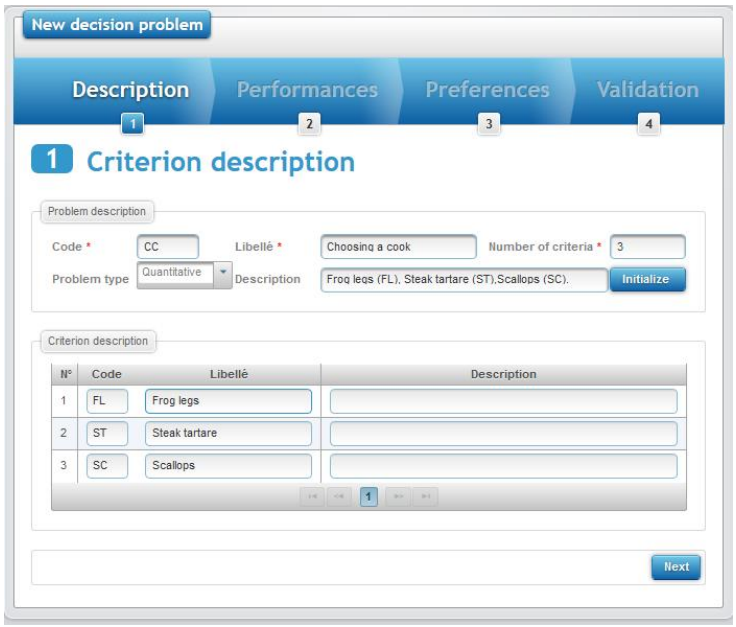


Figure 1: New decision problem

In this window, the first step allows to describe the new problem of decision support, as well as his different criteria. The second step, called Performance, allows us to enumerate the different alternatives and their evaluation for each criterion, called the performance matrix. The third step, called Preferences, allows you to define preferences between pairs of alternatives. And finally the fourth step and last, validates all of this information by saving them in the database.

After collecting data on the decision problem, the parameters of the chosen aggregation operator can compute. Depending on the type of problem (quantitative or qualitative), the choice of the aggregation operator will be based on the concerned category.

In the case of a decision problem, an aggregation operator is chosen in a category based on its effectiveness on similar decision problems. This is done by using a collaborative recommender system [8, 6] and by establishing a similarity model between decision problems. This similarity measure classifies decision problems into three broad categories depending on the nature of the problem. These are the questions of choice, sorting and storage, for more information [9]. The selected aggregation operator is tested by trying to determine its parameters from the preferences of the user. If the parameters of this operator

happen to be elicited respecting the set of preferences of the user, then it is proposed to the user, if not, another operator in the same category is chosen on the same bases. This procedure allows the user not to worry about the choice of the aggregation operator in the face of a decision problem and to obtain the best operator in the context of the use of the system.

A WEB PLATFORM RECOMMENDER SYSTEM:

We illustrate the system using following example. Four Chefs, a problem proposed by Marichal & Rubens [10]. We want to evaluate the chefs based on their ability to prepare three dishes: Frog legs (FL), Steak tartare (ST),Scallops (SC). The evaluation of the 4 chefs A, B, C, and D for each dish is given on a scale 0 to 20 in the following performance matrix:

Tableau 2: Evaluations of cooks

	FL	**ST**	**SC**
A	18	15	19
B	15	18	19
C	15	18	11
D	18	15	11

Reasoning of the decision maker:

- When a chef is known for his preparation of Scallops, it is better that he prepares Frog Legs well, as compared to Steak Tartare;
- Conversely, when a chef does not do a good job preparing Scallops, it is better that he prepares Steak Tartare well, as compared to Frog Legs.

Thus we can conclude than the decision-maker's ordering is A ≥ B ≥ C ≥ D.

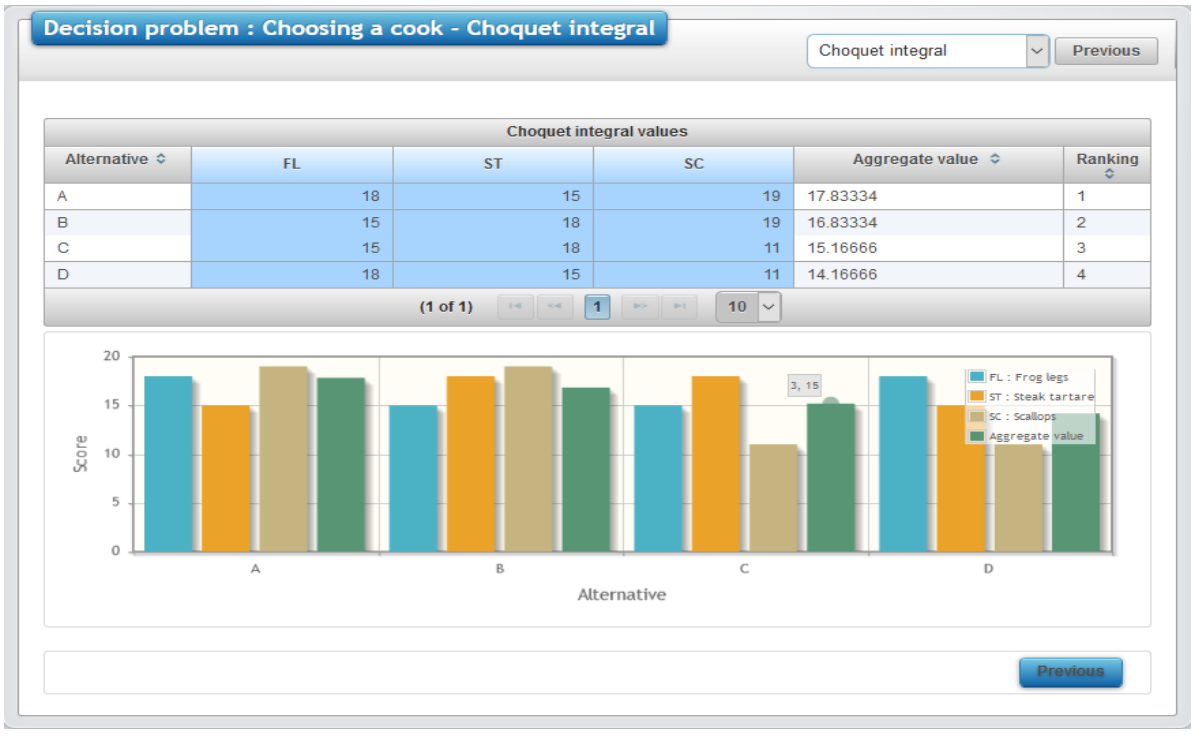


Figure 2:Results obtained by the recommender system on this decision problem

Results: one can easily check that the decision maker's preferences were taken into account.

CONCLUSION

Without being exhaustive, we presented some multicriteria aggregation operators used in problems of decision. We also have set up a generic recommender system whose choice of the aggregation operator is transparent for the user and is able to handle various problems of decision making, such as quantitative or qualitative problems. We use a collaborative model when choosing an aggregation operator in a decision support problem and a degree of satisfaction of the chosen operator. In the future, it will be reinforced by the integration of other aggregation operators and other decision-support concepts, such as the bi-capacity concepts [11]. It would also be interesting to propose new fuzzy measurement identification algorithms, faster and more robust, which tends to be greedy in time with a high number of criteria.

REFERENCES

[1] D. DUBOIS and H. FARGIER, "Making Discrete Sugeno Integrals More Discriminant," *International Journal of Approximate Reasoning,* vol. 50, pp. 880-898, 2009.

[2] D. DUBOIS, H. FARGIER and F. BONNEFON, "On the qualitative comparison of decisions having AAAI Press," *Journal of Artificial Intelligence Research,* vol. 32, pp. 385-417, 2008.

[3] G. CHOQUET, "Theory of capacities," *In Annales de l'Institut Fourier,* vol. 5, pp. 131-295, 1953.

[4] G. BELIAKOV, A. PRADERA and T. CALVO, "Aggregation Functions: A Guide for Practitioners.," *Studies Fuzziness and Soft Computing,* vol. 221, 2008.

[5] M. SUGENO, "Theory of fuzzy integrals and its applications. Thèse doctorale," Tokyo Institute of technology, 1974.

[6] S. FOMBA, P. ZARATE and M. KILGOUR, "A Recommender System Based on Multi-Criteria Aggregation," *IJDSST (to appear),* 2016.

[7] R. YAGER, "On ordered weighted averaging aggregation operators in multicriteria decisionmaking," *IEEE Trans. Systems, Man & Cybern,* vol. 18, pp. 183-190, 1988.

[8] K. LAKIOTAKI, N. F. MATSATSINIS and A. TSOUKIAS, "Multicriteria user modeling in recommender systems," *IEEE Intelligent Systems,* vol. 26(2), pp. 64-76, 2011.

[9] B. ROY, Méthodologie Multicritère d'Aide à la décision, Paris: ECONOMICA, 1985.

[10] M. GRABISCH, I. KOJADINOVIC and P. MEYER, "Kappalad : An R package for Choquet integral based MAUT," *63rd meeting of the EURO Working Group Multicriteria Aid for Decisions (MCDA),* 2006.

[11] J. MARICHAL and M. ROUBENS, "Determination of weights of interacting criteria from a reference set, European Journal of Operational Research," vol. 124, p. 641–650, 2000.

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

Assessing Supply Chain Risk through a Project Risk Management Approach

Mame Gningue,
Ecole de Management de Normandie
30, rue Richelieu – 76087 Le Havre, France
mgningue@em-normandie.fr,

Daouda Kamissoko
IMT Mines Albi-Carmaux
Campus Jarlard, 81013 Albi
daouda.kamissoko@mines-albi.fr

Sonia Froufe
Institut Supérieur d'Etudes Logistiques
Quai Frissard BP1137 Le Havre Cedex
sonia.froufe@univ-lehavre.fr

ABSTRACT

The intrinsic features of the Supply Chain have always put in the forefront the notion of uncertainty and risk in Supply Chain Management. However, Supply Chain Risk Management (SCRM) is a complex activity that compels to assess the vulnerability of all actors in the Supply Chain and the point would be to know which approaches of risk will allow to analyse and abate risks in the whole Supply Chain. Several approaches and methods have emerged in the literature to support organisations in their risk management. However, none appears to include all the commonly accepted stages in SCRM. To make up for this lack of integration, we have put forward in this paper an integrated approach for analysing the vulnerability of the Supply Chain inspired by Project Risk Management approaches. We detail an aggregative model of Supply Chain Vulnerability based on the aggregation of several criteria that could characterise it. Applying our proposal makes it possible to assess the viability of a Supply Chain in a specific context, in order to identify the most vulnerable actor in the chain and to put in place relevant risk reduction strategies.

Key words: Supply Chain, Risk Management, Project Management

INTRODUCTION

Our organisations are exposed to several types of risks and managing such risks have become a must for those managers who wish to remain successful in a highly competitive environment. Those risks must be assessed on the basis of their relevance and magnitude in order to identify vulnerabilities, draw up adequate plans to prepare for, and respond to, crises. In a Supply Chain, any risk that will materialise and affect one member or a process in the network, may interrupt the activities or the flows in several other members of the network. Supply Chain Risk Management (SCRM) is thus a very complex activity, in the sense the evaluation of the vulnerability of the Supply Chain as a whole takes into account that of the organisation concerned, but also that of the partners and the partners of partners.

SCRM approaches and methods have attracted much attention in the literature over the last two decades. All in all, four basic phases appear in the approaches detailed in research works: risk identification, risk evaluation, risk prioritisation, and risk processing. Although there is a significant relationship between these various phases, authors have generally focused on two of these phases. To compensate for this lack of integration, we put forward in this paper an integrated approach for analysing Supply Chain vulnerability by drawing from Project Risk Management approaches.We detail an aggregation model for Supply Chain vulnerability based on several criteria that can characterise it. The implementation of our proposal allows to judge the viability of a Supply Chain in a specific context, in order to identify the most vulnerable actor in the chain and to put in place risk reduction strategies. We have applied our SCRM model to a real- life case study to test the feasibility and efficiency of the model suggested. This implementation has allowed us to identify the level of vulnerability of each and every actor in a Supply Chain, to draw the attention of all members on which actions should be engaged in to reduce the vulnerability of the entire chain.

SUPPLY CHAIN RISK MANAGEMENT

The Supply Chain is a network of organisations that take part in the various upstream or downstream processes and activities which create value in the form of products and services brought to the consumer [1]. Managing the Supply Chain means managing upstream and downstream relationships with suppliers and customers, at a lower cost throughout the entire Supply Chain [2].

The intrinsic characteristics of the Supply Chain (multi-actors, multi-sites, multi-flows) have always brought to the forefront the notion of uncertainty and risk in Supply Chain Management. Recent events across the world and their consequences have confirmed the vulnerability of Supply Chain and that it is no longer possible to reach a certain level of performance without integrating risk management into the management of global Supply Chains. This what gave birth over the last decades to Supply Chain Risk Management (SCRM).

In their work, Sodhi, Son, and Tang [3] have found out that there is a lack of consensus on the definition of the concept of SCRM. Lindroth and Norrman [4] consider that SCRM lies where Supply Chain partners collaborate and apply risk management process

tools to manage and alleviate uncertainty and risks caused by, or bearing on, activities and resources linked to the Supply Chain. Juttner et al. [5] define SCRM as "the identification and management of risks for the supply chain, through a co-ordinated approach amongst supply chain members, to reduce supply chain vulnerability as a whole." However, on the whole, authors in the literature agree on four basic phases in SCRM [5-10]: risk identification (to identify what could happen, where, when and how), risk evaluation (to quantify/measure the impact risks may have), risk prioritisation (to prioritise those risks identified), risk processing (to develop and implement an attenuation strategy to control risks).

SCRM approaches have also been the subject of much research work but research works that put forward SCRM processes have on the whole studied each of the phases in the process individually [11]. Indeed, researchers who presented integrated processes have mainly focused on two phases in the SCRM [11]. The risk attenuation phase has attracted more attention, with assessment coming second. Both these phases are largely dealt with through quantitative methods inasmuch as risk assessment includes the quantification of probabilities of occurrence and risk event impacts. Those qualitative approaches used in integrated SCRM processes are conceptual by nature or merely explain SCRM steps or phases but do not demonstrate how such approaches should be applied. Since there is a significant link between all phases in SCRM, more attention should be legitimately given to integrated processes in general.

Moreover, not only does the literature lack adequate processes to respond to Supply Chain crisis situations, but there is a lack of empirical studies on Supply Chain risk analysis as well [3].

SUPPLY CHAIN MANAGEMENT AND PROJECT MANAGEMENT

There are forceful similarities between the Project concept and that of Supply Chain. The project is defined as 'a single process which consists in a set of coordinated and controlled activities that comprises starting and ending dates, undertaken with a view to reaching an objective complying with specific demands, including deadline, cost and resource constraints' [12-p149]. Its goal is to produce a single result in the form of value to the customer. Project management is equally a structured process for managing work flows in a specific order spreading over time, with a starting and ending date, and specific entries and results.

When it comes to project management, mastering those risks linked to the specifications is a major concern. Thus, several activities are implemented iteratively throughout the duration of the project to manage any risk which might threaten the success of the project. Those activities will generally be project risk analysis, risk reduction, and monitoring. Putu et al. [13] became interested very early in the similarities which exist between project management and Supply Chain management. These authors consider that these two modes of management have similar features in the sense that they both operate in a complex, dynamic and uncertain environment [13]. Furthermore, some research works have recently underlined how close temporary multi-organisational project risks and Supply Chain management were [14].

By drawing more widely from the works by Froman and Gourdon [12] and by Mentzer et al. [15] on Supply Chain Management, we are able to build an analogy between both these two concepts, an analogy that is summarised in the table below.

	Project [12]	**Supply Chain [15].**
Structure	Set of activities	Set of processes
Actors	Different actors and organizations	Different actors and organizations
Objective	Comply with specific customer needs	Satisfy customers demand
Constraints	Delay – Cost – Resource	Delay – Cost – Capacity

These similarities between project management and Supply Chain management, together with the closeness between types of risks in both these domains, have prompted us to draw our inspiration from project risk management approaches to develop a model for SCRM.

SUPPLY CHAIN RISK MANAGEMENT MODEL

In this section, we wish to put forward an approach for assessing Supply Chain vulnerability with respect to Risks. Risk is used here in the meaning of feared event that can disturb one or several activities of the supply chain. We assume that the latter Supply Chain is already in operation. Our purpose is not to look into the design of a new Supply Chain.

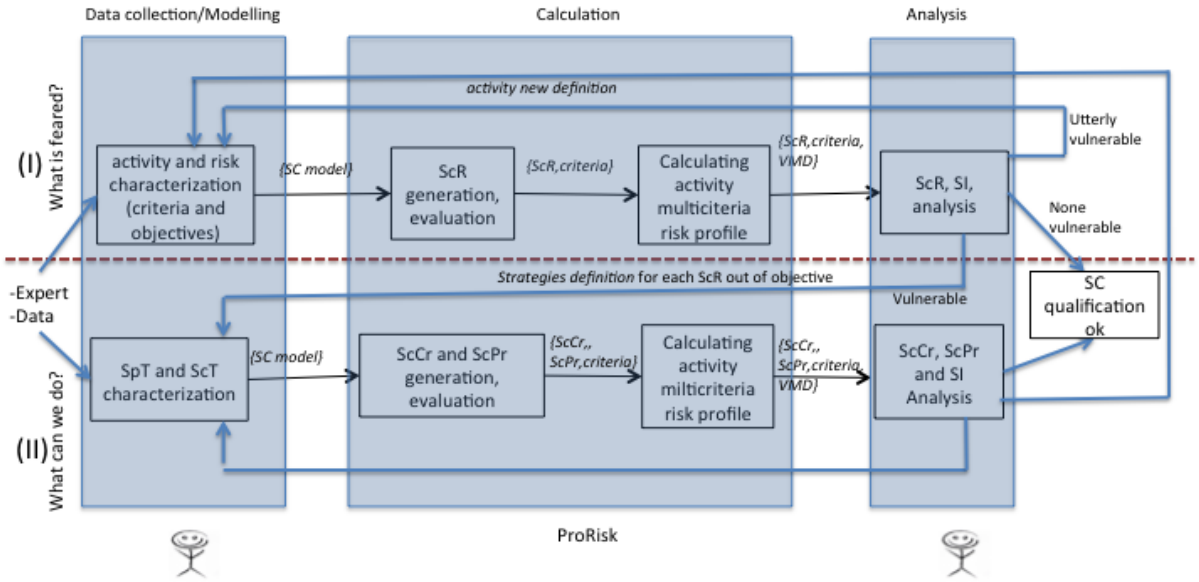


Figure 1: Risk management model

The main question we wish to address thanks to an integrated SCRM model is the following: Given a set of risks and set objectives on several criteria, is the Supply Chain vulnerable? The finality of this question is to determine the combination of activities and/or providers that failure lead to the supply chain vulnerability. The approach we adopt to answer this question will begin with an identification of the range of feared events at the level of each activity. Every activity being characterized by several criteria including cost and duration. Then, we will assess risk factors that allow to characterise these situations, thus enabling us to classify not only risk factors in order of importance but also entities as regards their level of vulnerability. Finally, we will suggest a variety

of actions likely to be activated in order to reduce globally the vulnerability of the Supply Chain. The following diagramme presents our recommendations in this respect:

On the whole, our approach may be summarised in three main phases: data collection and modeling, calculation or assessment, and analysis. Each phase comprises two possible actions in accordance with the analysis situation.

Data collection and modeling are to be achieved by an analyst. They consist in identifying Supply Chain activities, potential risks, and stakeholders' objectives. Each activity is characterised by several criteria (e.g., cost). Sequencing links and implementation conditions are identified between activities. Risk occurs with a certain frequency and may impact one or several activities. Objectives are also defined with respect to several criteria (e.g., cost objectives). By the end of this phase, we will produce a Supply Chain model. This model is a representation of the previous data. Then, we will be in a position to define possible risk scenarios (ScR). A risk scenario is a combination of several risks. It includes also single risks and the situation where no risk occurs. The Supply Chain will thus be assessed for each pair (ScR, Criterion). A multidimensional vulnerability indicator (VMD) will allow to characterise the Supply Chain on the basis of the pairs (ScR, Criterion). The multidimensional aspect of vulnerability is imposed by the fact several criteria are to be taken into account. Each criterion is a dimension. Our risk management tool, ProRisk, supports the generation of scenarios and the determination of vulnerability according to the vulnerability model. The VMD analysis allows to classify SC in three categories: non-vulnerable, vulnerable and utterly vulnerable. If the Supply Chain is non-vulnerable, the situation is acceptable and no action should be contemplated. The SC can continue its activities. There is any risk that can disturb consequently its good functioning. Utterly vulnerable means that the Supply Chain does not meet objectives in any way. The gap is deemed unacceptable. In this situation, some risks scenarios can disturb consequently the good function of the supply chain when they occur. In this case, we characterise again activities, or even objectives. The characterization means determining a new value of their attributes. The intermediate situation is one whereby the Supply Chain is viewed as vulnerable. To reduce this vulnerability, strategies are devised that can be applied preventively and/or correctively. A strategy is defined as a set of actions aiming at adding a new activity, to modify or cancel it. They aim at dealing with a specific risk by reducing its impact. Preventive strategies may possibly modify the probability of an occurrence of a risk for which they have been contemplated. Added activities are to be characterised in the same way as initial activities. ProRisk then generates scenarios for corrective or preventive treatment (ScCr et ScPr) for those risk scenarios which did not comply with the objectives. The new value of the vulnerability analysed will classify the Supply Chain in one of the three previous categories. This steps are repeated as many times as necessary until an acceptable level of risk is reached.

This proposal is useful for the manager of the supply chain. In fact, in the situation of none acceptable vulnerability, it become possible to identify all possible risk scenario. Thus, for every scenario, we can identify the involved activities and providers. To validate the model is has been applied to the supply chain of a food bank.

CONCLUSION

Disruptions in the Supply Chain may impact on organisations and their partners. They may lead to useless reactions and interventions, assumptions, lack of trust between partners, higher costs and lower efficiency. They may even go as far as to impair the financial performance of the chain as a whole. It is thus critical to protect Supply Chains and to get vulnerability reduction strategies exercised by actors.

Opportunities for future research in SCRM as a domain are manifold. Applying decision-making models in SCRM is a topic for research to be explored. It is also necessary to assess the performance of SCRM models with primary data, as well as their adaptability and flexibility by applying them to different companies in identical or different sectors.

The potential analogy between project management and Supply Chain management directed us towards applying project risk management approaches to SCRM. To meet the need for integrating the various phases in SCRM, we have suggested a model based on an approach that spread from risk identification and likely risk scenarios to implementing reduction strategies through assessment of the vulnerability of actors and the Supply Chian as a whole. In a coming phase, this model will be applied to an existing Supply Chain whose vulnerability will be assessed as and when attenuation strategies will be applied between actors.

REFERENCES

[1] M. Christopher, *Logistics and Supply Chain Management: Strategies for Reducing Costs and Improving Service*. New Jersey: Financial Times Prentice Hall, 1992.

[2] M. Christopher, « Managing risk in the supply chain », *Supply Chain Practice*, vol. 7, nº 2, p. 4–21, 2005.

[3] M. S. Sodhi, B.-G. Son, et C. S. Tang, « Researchers' perspectives on supply chain risk management », *Production and Operations Management*, vol. 21, nº 1, p. 1–13, 2012.

[4] R. Lindroth et A. Norrman, « Supply Chain Risk Management: Purchasers' Vs. Planners View on Sharing Capacity Investment Risks in the Telecom Industry. », in *Proceedings of the IPSERA 11th International Conference*, 2001, p. 25–27.

[5] U. Jüttner, H. Peck, et M. Christopher, « Supply chain risk management: outlining an agenda for future research », *International Journal of Logistics: Research and Applications*, vol. 6, nº 4, p. 197–210, 2003.

[6] C. Harland, R. Brenchley, et H. Walker, « Risk in supply networks », *Journal of Purchasing and Supply Management*, vol. 9, nº 2, p. 51-62, 2003.

[7] L. M. Hauser, « Risk-adjusted supply chain management », *Supply Chain Management Review*, vol. 7, nº 6, p. 64-71, 2003.

[8] P. R. Sinha, L. E. Whitman, et D. Malzahn, « Methodology to mitigate supplier risk in an aerospace supply chain », *Supply Chain Management: an international journal*, vol. 9, nº 2, p. 154–168, 2004.

[9] J. Hallikas, I. Karvonen, U. Pulkkinen, V.-M. Virolainen, et M. Tuominen, « Risk management processes in supplier networks », *International Journal of Production Economics*, vol. 90, nº 1, p. 47–58, 2004.

[10] P. R. Kleindorfer et G. H. Saad, « Managing disruption risks in supply chains », *Production and operations management*, vol. 14, nº 1, p. 53-68, 2005.

[11] W. Ho, T. Zheng, H. Yildiz, et S. Talluri, « Supply chain risk management: a literature review », *International Journal of Production Research*, vol. 53, nº 16, p. 5031–5069, 2015.

[12] B. Froman et C. Gourdon, *Dictionnaire de la qualité*. Saint-Denis-la-Plaine: AFNOR, 2003.

[13] P. D. Karningsih, B. Kayis, et S. Kara, « Development of knowledge based systems for supply chain risk identification in multi-site & multi-partners global manufacturing supply chain », in *The 13th Asia Pacific Management Conference, Melbourne, Australia*, 2007, p. 466–471.

[14] A. M. T. Thomé, L. F. Scavarda, A. Scavarda, et F. E. S. de S. Thomé, « Similarities and contrasts of complexity, uncertainty, risks, and resilience in supply chains and temporary multi-organization projects », *International Journal of Project Management*, vol. 34, nº 7, p. 1328-1346, oct. 2016

[15] J. T. Mentzer *et al.*, « Defining supply chain management », *Journal of Business Logistics*, vol. 22, nº 2, p. 1-25, 2001.

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

Optimization to support policy decisions stimulating value creation in the forestry sector

Ulf Johansen, Adrian Werner, and Vibeke Nørstebø
SINTEF Technology and Society
7465 Trondheim, Norway
Ulf.Johansen@sintef.no, AdrianTobias.Werner@sintef.no, Vibeke.S.Norstebo@sintef.no
web-page: www.sintef.no

ABSTRACT

Following past decades' extensive afforestation efforts in Norway, mature forest volumes are currently increasing. Forestry politics call for sustainable and efficient resource usage and increased regional processing, improving regional value creation. We demonstrate how a blend of methods from operations research and regional macro-economics may be utilized to support policy and decision makers in this process. The operations research perspective is concerned with finding an optimally designed wood value chain and an aggregated planning of its operations, taking a holistic perspective on strategic-tactical level. Using Input-Output analysis methods based on statistics and survey data, regional macro-economics helps to estimate the impact on society beyond immediate value chain activities. Combining these approaches in a common mathematical optimization model, a balance can be struck between business and regional political interests. We demonstrate the concept on a stylized case investigating the effects of road transport subsidies on efficient resource usage, transportation, industry development and value creation. Concluding, we discuss how the found insights may support public administration to explore policy effects more comprehensively and to form effective measures stimulating growth in rural areas.

Keywords: DSS in Forest Industry, Value Chain Analysis, Multi-Criteria Optimization, Regional Economy, Resource Efficiency

INTRODUCTION

Historically, forestry has been an important industry in Norway, both for value creation and for employment. Obviously, this importance declined with the rise of new industries. Now, the Green shift creates large opportunities again for the forest sector and forest-based industries to play a vital role in the global bioeconomy. The European Commission's Bioeconomy Strategy [1] highlights the crucial role of forests in the global carbon cycle and the fight against climate change. Growing demand for wood products provides a strong stimulus for increasing forest resources and productivity. This calls for improved management practices as well as knowledge and tools to help supplying bio-based raw

materials in a resource efficient way. At the same time, incentives and policies are needed to foster thriving rural livelihoods [1].

In a European context, Norway is the country with the largest share of unexploited forest resources [2]. District and regional policies aim at strengthening conditions for natural-resource based industries [3], ensuring the continued existence of rural employers. The wood sector's value creation is set to quadruple by 2045, compared to 2012 values [4], for example, through more regional wood conversion and processing activities and more efficient resource usage. Much of the anticipated increase in value creation is set to happen in rural areas, and coastal regions such as northern Norway are expected to play a central role in this regard.

However, current logging activities are outpaced by the growth of mature forest resources. The annual growth in forest volume in coastal Norway increased from 2.8 million m^3/year in 1925 to 8.5 million m^3/year in 2015, leading to a standing stock in 2015 of 300 million m^3 with a sales potential of 60 billion NOK [5]. In contrast, only 2 million m^3 were harvested in 2015 [5].

Questions of how to increase logging activities, supporting regional timber processing and stimulating regional value creation, motivated the development of a business-economic optimization model taking into account regional economic effects [6]. With a holistic view on the wood value chain (VC), it constitutes one of few approaches combining value chain optimization and regional macro-economic theory by way of Input-Output (I-O) analysis methods. The model has been developed with partners from industry and public administration in coastal northern Norway, intending, among other goals, to support local and regional policy and decision makers. Although a mathematical model never can capture all reality's complexities, it can contribute to explore effects of various policy instruments in an easy and comprehensive way. Bearing in mind that the VC actors respond to policies by re-optimizing their operations, it helps to point out policy impacts along the whole value chain, highlighting also unexpected effects. Optimizing under different objectives, the situation can be probed from various viewpoints, balancing regional politics, industry interest and environment protection. The model reflects regional industry structures, forest resources and infrastructure, ensuring applicability and local relevance of the found insights.

A report to the Stortinget [7] highlights the need for sustainable use of forest resources, for efficient transport and for a competitive process industry. Among suggested measures are further infrastructure development and balancing of public policies and private interests. Relevant public policy instruments currently applicable to the forestry and related transportation sectors are outlined in the lower part of figure 1 in black and red, respectively. The majority of these instruments aims at the primary segment of the value chain, logging activities [8]. Only the regionally differentiated payroll tax (RDP) system applied in Norway, subsidising labour costs in rural areas, may reduce costs for several actors in the value chain. Subsidies on transport costs in rural areas as presently found in, e.g., Sweden [9] or Finland [10] constitute an alternative policy with similar aims - stimulating growth in rural areas.

In this paper, we demonstrate the concept of our optimization model and how it may support regional authorities' decision making by way of a stylized case study. As our model does not address labour costs explicitly, we consider a hypothetical situation of introducing road transport subsidies in a Norwegian region. Following an outline of our optimization model and the considered case, we explore results of a simplified policy analysis. Concluding, we discuss how the model can support authorities in finding efficient policy instruments stimulating regional resource usage and business development.

THE DECISION SUPPORT MODEL AND CASE STUDY

We consider a value chain tracing the flow of raw and processed forest products as outlined in figure 1. Important elements are forest owners / timber traders, various companies processing wood products, and end-user customers, connected by transportation links including harbours. Transport happens mainly by truck or ship. The optimization model takes a strategic-tactical perspective and addresses harvesting plans, transportation, production of wood and timber products as well as industry development in combination with economic ripple effects (REs) of these activities on the wider society. It seeks to optimize the profit within the VC, national and regional REs from logging and processing or a combination of both. Revenues accrue from forest owners' sales of timber and pulpwood and from companies' sales of products to other companies and end-users. Costs arise from logging, from transportation and storage, as production costs including purchase of timber and other production factors and from investments into new production lines or companies. To estimate REs, we use I-O multipliers determined for each actor in the VC, relating to their (optimized) production levels. Pre-calculated from statistics complemented by company surveys, the multipliers reflect the actual industry structure in the studied region. The model is a mixed-integer linear programming (MILP) problem, implemented in the FICO™ Mosel modelling language. Typical problem instances contain about 280.000 constraints and 1.4 million variables and are solved by standard solvers in five to ten minutes.

For a policy analysis, one may investigate characteristics such as profits achieved for the whole VC or for selected groups of participants (e.g., regional forest owners, regional companies), created REs, usage of regional resources or investment behaviour and compare to a baseline solution reflecting the current state. We apply the model to a case region in coastal northern Norway with known mature forest resources. In addition to a replication of the existing industry structure with sawmills, particle board production and pulpwood processing, we include hypothetical options to establish sawmills or pulpwood processing industry. In a baseline scenario, all transportation costs correspond to values reported in the literature or statistics. We investigate how an introduction of road transport subsidies would affect regional resource usage, production and investment decisions across the value chain and, consequently, profits and value creation (REs) over a medium-term period (25 years). Transport subsidies were earlier available in the area and are in use in adjoining Swedish regions. Therefore, they are an often discussed policy instrument, not only with respect to potential distortion of competition. The Swedish subsidy rate is differentiated by location and transportation distance, varying between 5% and 45% [9]. In a simplified analysis, we study a flat rate of 20% subsidy on all distance dependent road transport costs.

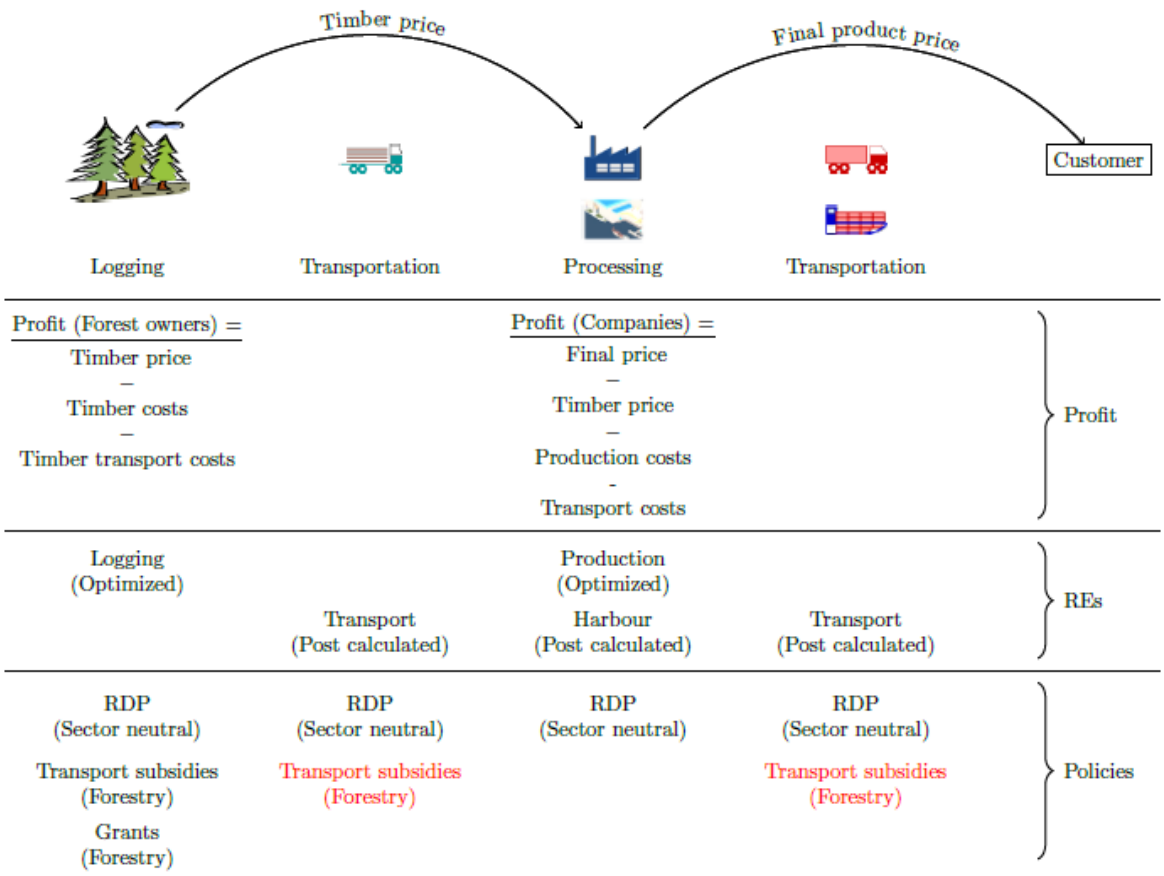


Figure 1: Activities included in VC profit, RE estimations and current policies.

POLICY ANALYSIS RESULTS AND DISCUSSION

The aim of subsidising road transport costs is to stimulate growth in rural areas. One should expect that this leads to increased road transport due to decreased costs and higher profitability margin for the forest industry. It should also result in increased harvesting at locations further away from industry, stimulating an expansion of processing industry.

Results from the case study are summarized in table 1. The second row indicates how the subsidy affects the level of distance dependent road transport costs for the whole VC. Compared to the baseline scenario, they are reduced by 19%, almost equal to the subsidy rate of 20%. Note that this reduction by 262 MNOK also involves differences in the total road transportation distance, which increased by 1.5% from 90 645 000 km to 91 977 000 km. Corrected for this, the transport subsidy paid by the authorities was 282 MNOK. We investigate now whether the effects from this subsidy are congruent with authority ambitions.

Table 1: Results for baseline and subsidy scenarios (all economic values in MNOK).

	Baseline	Subsidies	Δ	Δ %
Total road transport costs	4 069	3 770	-299	-7%
Distance dependent costs	1 391	1 129	-262	-19%
Other costs and fees	2 678	2 641	-37	-1%
Total VC profit	17 588	17 794	206	1.2%
Companies' profit	11 393	11 469	76	0.7%
Company investments	94	239	145	155%
Forest owners' and timber traders' profit	6 195	6 325	129	2.1%
New processing industry (# investments)	1	3	2	–
Use of regional forest resources	61%	59%	–	–
National economic ripple effects	19 456	19 254	-202	-1%
Regional economic ripple effects	5 573	5 466	-107	-2%

The transport subsidy improves VC profit by 206 MNOK in total, whereof 76 MNOK for companies and 145 for forest owners / timber traders. Also, investment activity increased: In the baseline scenario, production capacity is expanded by one pulpwood processing company, while it is optimal to invest in two more companies when the subsidy is introduced. These results comply with policy goals, but the VC profit improvement is 76 MNOK lower than the total transport subsidies paid out. An immediate explanation is that the additional two companies incur extra investment expenses of 145 MNOK, which clearly dampens the total VC profit level.

However, harvesting and use of regional forest resources decreased from 61% to 59% despite cheaper road transport and the establishment of new pulpwood processing companies. The total production of the three new companies is slightly lower than the production of the new company in the baseline scenario while production levels of the existing industry remain unchanged. This means that increased investment activity does not necessarily implicate higher total production – a rather unexpected result, which may be explained by better resource efficiency due to better spatial distribution of the added production capacity.

The increased road transport distance is accompanied by a decrease in ship transport distance. From an environmental perspective, this may not be a desired effect and may not correspond to public authority goals. An effect of these changes in road and ship transportation is a spatial redistribution of harvesting between the region's municipalities: Harvesting activities reduce mainly in coastal municipalities or on islands in the northernmost part, furthest away from large, resource-demanding industry. They increase in an inland municipality closer to new and existing companies.

Although use of regional forest resources reduced, transport subsidies do foster rural development due to more diversified investments. In the model, economic REs are assumed proportional to production output. This is a simplification, and more new and distributed companies may create additional local REs such as rural job creation and corresponding consumer effects. Also, the three new companies have a higher total production capacity, increasing the sector's robustness and flexibility to react to possibly growing demand for new and existing forest products.

CONCLUSIONS

As a result of past decades' extensive afforestation, mature forest volumes in Norway are increasing. National forestry politics calls for sustainable and efficient resource usage and increased regional processing. Against this background, we developed a dynamic optimization model combining VC activities and aspects of regional macro-economics. The model considers the wood VC from logging to sales of end products. Reflecting regional resources and relationships, it ensures applicability and local relevance of the found insights. To exemplify the usage of the model for decision support, we apply it to a stylised policy analysis case, comparing a baseline scenario reflecting the current situation to a scenario with road transport subsidies. The case illustrates that the subsidies lead, as expected, to an increase in road transport and VC profits. They also contribute to regional business development as it becomes more profitable to establish new processing industry. However, other effects are a redistribution and slight decrease of logging activity and a decrease of ship transport. While the former effects appear to be in line with the subsidy policy goals, the latter contrasts official goals. Most likely, this would have gone undetected without a more holistic assessment of wood VC interrelations, beyond the immediate goals of the policy, and a view of the VC as a complex network of actors affecting each others' decisions simultaneously.

Compared to traditional approaches to analysing industry, economic and societal effects of regional policy, a more comprehensive VC optimization can help to unearth deeper insights. For instance, our analysis illustrates that a transport subsidy may have wider or different effects than anticipated. While our model focuses on economic effects in the wood VC, follow-up analyses may investigate consequences for other industry sectors, emissions from transport and industry, needs for (or avoidance of) infrastructure upgrades and connected environment interventions, land use or other effects on local society. In a longer-term perspective, it may be worthwhile to include further sustainability aspects in a VC optimization model. Given the complexity of many ecosystem service models, a full integration may prove challenging. Impact factors and relations, which – like economic REs – can be described by rather simple functions may be a good start.

Optimization models such as the one discussed here simplify a complex reality but contribute to exploring effects of policy variants in a rather easy and comprehensive way, with respect to counter-intuitive mechanisms. Allowing to investigate a situation from several viewpoints, interests of various involved parties can be explored to find a balance between regional politics, industry, societal interests or environment protection. To support regional policy finding, the model can be tailored to the characteristics of the considered area, exploiting local expertise. Such models have the potential to bring authorities and industry together on a common ground to understand the wider implications of policy making. In this manner, they complement and support simulations and more qualitative policy assessments. This larger picture should always be kept in mind, not at last since quantitative decision support approaches typically show a simplified or idealized situation and need to be set in a wider context.

ACKNOWLEDGMENTS

This work is part of the NorFor project funded by the Norwegian Research Council (NRC grant no. 239093) and by partners from industry, business forums and public administration. The research has also been supported by the RegModell project (NRC grant no. 102001007). We thank the reviewers for their helpful comments to improve this paper.

REFERENCES

1. European Commission. Innovating for sustainable growth – A bioeconomy for Europe, 2012.

2. A.-M. Remøy, A. Sundby, B. A. Øveraas, K. Bruvoll, L. Fjørtoft, M. Valseth, O. J. Marvik, S. Thormodsdottir, and Ø. Halvorsen. Bioøkonomien – et viktig bidrag til fremtidens vekst og verdiskaping. (In Norwegian), Innovasjon Norge, 2015.

3. Kommunal- og moderniseringsdepartementet (KMD). Ta heile Noreg i bruk. Meld. st. 13 (In Norwegian), 2012.

4. Landbruks- og matdepartementet (LMD). SKOG22. Nasjonal strategi for skog- og trenæringen. Official report (In Norwegian), 2015.

5. Kystskogbruket. Melding om Kystskogbruket 2015. Official report (in Norwegian), 2015.

6. U. Johansen, A. Werner, and V. Nørstebø. Multi-criteria optimization of the forestry value chain in northern Norway taking into account national and regional economic trade-offs. Submitted, 2017.

7. Landbruks- og matdepartementet (LMD). Verdier i vekst – konkurransedyktig skog- og trenæring. Meld. St. 6 (2016-2017) (In Norwegian), 2016.

8. E. Bergseng and B. Solberg. Evaluating financing of forestry in Europe (EFFE). Country level report – Norway. Technical Report INA fagrapport 8, Institutt for naturforvaltning, Universitetet for miljø- og biovitenskap, 2007.

9. Tillväxtverket. Transportbidrag til Norra Sverige (In Swedish). https://tillvaxtverket.se/aktuella-amnen/starta-foretag/stod-till-foretag/transportbidrag-till-norra-sverige.html.

10. Trafikanalys. Transportpolitiskt stöd till perifera regioner i Sveriges grannländer (In Swedish). Technical Report 2015:9, 2015.

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

Using a negotiation support system in language teaching

Margaret Kersten
SLaLS, Carleton University, Ottawa, Canada
E-mail: margaret.kersten@carleton.ca

ABSTRACT

International and immigrant students at Canadian universities come from culturally, linguistically, and educationally diverse backgrounds. Those who do not meet an English language requirement are enrolled in English for Academic Purposes (EAP) courses. These courses focus on the development of academic language skills and critical thinking. In this paper we show that there is a significant impact on achieving these courses' objectives by using Inspire — an online negotiation support system. An average increase in the final assignment grade of 5% was found when using Inspire versus traditional materials in an advanced level EAP course.

Keywords: Negotiations, online negotiations, teaching, experiments, EAP.

INTRODUCTION

The use of computer technology in the language classroom has a long history, from early *behaviourist* computer assisted language learning (CALL) programs in the 1960's, through *communicative* CALL with programs such as "Sim City" and "Where in the World is Carmen Sandiego" in 1970's and 1980's to *integrative* CALL in the 1990's, where the most popular applications have been email, electronic boards and chat rooms. These early programs were criticized for being restrictive or for being used in an ad hoc manner, with no clear pedagogical focus or sense of direction [1]. The same criticism may apply to the use of social media in today's language class. Unless the learning objectives are clearly formulated and executed, the outcomes may be meagre, particularly in teaching English for Academic Purposes (EAP) at a university level.

More recently developed language learning software such as "Rosetta Stone", "Reading Assistant", "Reading Coach" have been criticized for removing learners from an authentic learning environment with the language activities being out of context. Instead of engaging students in social interaction in an authentic setting, students practice listening, speaking, reading, and writing as separate areas [2]. This contradicts language acquisition theories [3] which emphasize language's social and communicative functions [4]; promote "natural" acquisition of language from texts, peers and teachers; emphasize the role of "authentic" texts (not simplified for language learners); and put the student in the center of the learning process. Instructional models promote spending an extended period of class time on one theme, where the four language skills are integrated. Collaboration among students on problem solving as well as sharing ideas and perspectives promotes social development and helps students progress from

a lower to a higher level of language skills and knowledge [5].

The technology that has been most widely used for collaboration between different groups of students in different locations is email. Email lends itself to collaborative learning projects. However, because email is just an electronic platform, the instructors designing collaborative projects are not supported in any way. They are responsible for creating groups, matching students, structuring learning activities, and monitoring group interaction [6]. A system that takes over some of those responsibilities would benefit not only the teachers but also the students who often have difficulty following oral and even written instructions.

A quick internet search shows that the use of DSS in English for Academic Purposes (EAP) courses is very limited. There are only a few knowledge-based systems, for example, ET-DES: English Tense Diagnosis Expert System [7]. This system draws learners' attention to their errors in using English tenses, but it does not contribute to the development of communicative competence; there is no social context, no meaningful interaction, no communication.

When designing courses, EAP teachers draw on sources from different academic fields, so a knowledge-based system that supports professional training can be adapted to EAP. However, doing so would not meet the objective of developing communicative competence in a social context. Systems that do meet this requirement are negotiation support systems (NSS), particularly those that have been used in management training [8],[9],[10].

NSS use analytic tools that help users solve problems and make decisions. One of their applications is teaching and training of professionals in different fields. In language teaching, the benefits of implementing problem solving have been widely discussed. For example, Buerke [11] argues that problem-solving is meaningful as it relies on students' natural tendency to figure things out, which enhances language acquisition. Norman and Schmidt [12] maintain that problem based learning contributes to intrinsic motivation and develops self-directed learning skills, which tend to be maintained. Thomas [13] showed that students in schools which implemented problem-based learning demonstrated significant improvement on standardized tests of academic achievement. These findings suggest that adoption of an online negotiation system for language teaching should lead to positive learning outcomes.

This paper reports on the results obtained from the comparison of students' grades in two sections of an advanced level EAP course. One section used only traditional materials throughout the course; the other used online negotiations that substituted one of the course components.

ENGLISH FOR ACADEMIC PURPOSES

International and immigrant students who wish to pursue their academic career at Carleton University, where the author works, are required to take one of the internationally recognized English language proficiency tests. Those who do not meet language requirement are placed in EAP courses.

The EAP courses at Carleton University are designed using the principles of sustained-content instruction. Sustained-content courses focus on one theme and are therefore similar to "regular" university courses. It has been argued that exploration of one subject area over a longer period of time builds content knowledge and lends itself to the development of critical thinking [14]. The instructional model is embedded in the theory of second language acquisition, which makes a distinction between the "natural" way of "picking-up" language and conscious learning [3]. The theory posits that acquisition plays a central role in developing linguistic competence and learning is secondary. For acquisition to occur, learners need comprehensible input from texts, peers, and teachers.

The EAP students come from diverse linguistic, cultural and educational backgrounds. Moreover, their academic and professional interests vary widely. Therefore, an effort is made to identify course themes that are relevant to all students, regardless of their field of study.

A typical advanced EAP course has the following components: 1) a course pack with readings from literature, 2) a research project and 3) a module that focuses on analyzing cases. The three components are thematically interrelated and build on each other in order help the learners to move from a lower to a higher level of content knowledge and language proficiency.

For the purpose of this study two sections of an advanced EAP course were selected. one used the theme of "diffusion of innovations" and the other "decision making and negotiations". The course objectives and the structure were the same. However, the content of the third component differed – one section used readings from theory of innovations diffusion to analyze cases; the other section used online negotiations to hone students' analytical skills. The NSS that was used supported the students in finding a solution to their negotiation problem. At the end of the third module both groups wrote a report.

Based on findings from literature on the benefits of implementing problem solving in language teaching [5,11,12,13], the following hypothesis was formulated:

Hypothesis: Replacing one component of a traditional course with the online negotiation component increases language proficiency as measured with students' grade increases.

In order to determine if the adoption of Inspire resulted in better learning outcomes we compared grades in two sections of the same course, one traditional (Traditional) and one with online negotiations (Inspire). The objectives, assignments, structure and teaching methodology were the same. In both the final in-class assignment was a report. The sections were taught by the same instructor and the students' backgrounds were similar.

THE INSPIRE COMPONENT

The course that used Inspire as a tool for language learning focused on the topic of decision making and negotiations. As mentioned earlier, EAP students have different academic interests but this topic is relevant to virtually any field of study. There is another aspect of decision making and negotiations that renders it suitable for EAP courses. Making decisions and negotiating are ubiquitous human activities; everybody makes decisions and everybody negotiates. This allows teachers to tap into students' background knowledge or schemata. Schema theory explains that background knowledge learned in the first language helps learners to understand texts in the second language [9]. To that end, adoption of an NSS system satisfied the need to create opportunities for the learners to naturally acquire language.

The system used for the course was Inspire. Inspire integrates computer-based technologies with decision and negotiation support tools and communication platform that support negotiators throughout the negotiation process [10]. The negotiations are anonymous which allows negotiators to focus on issues rather than personalities. They are also asynchronous, giving negotiators time to carefully analyze and consider options using the system's analytical tools.

Inspire is based on the phase model of negotiations. In the pre-negotiation phase the students prepare for negotiations: they fill in a pre-negotiation questionnaire, watch a demo, read the case and rate issues and options. Then they formulate their first offer and a message to their counterpart. During the negotiations, the system provides visual tools depicting the history of negotiations, which help participants analyze their progress and make decisions. The negotiations end when an agreement is reached or when the deadline expires. Then the participants can enter the post-negotiation phase to look for other solutions. Finally, they fill out a post-negotiation questionnaire.

Inspire is easy to use, the case involves a contract negotiation between an agent representing a young song writer and singer and a mid-level manager of an entertainment agency. The process is highly structured and requires that students follow instructions closely and pay attention to the system's features. This allows them to think critically and analyze theirs and their counterpart's moves and behaviour. In using the system's support tools, the students can make informed decisions regarding the strategy and the message that they plan to send to their counterparts. Student that partook in the negotiations were told that the outcome did not matter; what mattered was the process — that is, serious engagement in the negotiation. The number of offers and email exchanges ranged from four to sixteen, with the majority of students exchanging between five and seven offers. When the negotiations ended, the participants filled out a questionnaire in which they evaluated their counterpart, themselves, and the process.

Even though Inspire is easy to use, the process was not without glitches; 50% of the group had to be reminded a few times to register before the registration deadline, a few forgot their user name and password; 25% of the group had to be reminded to negotiate; one student had an unresponsive counterpart and did not report this to Inspire team for three days, which gave them very little time to assign a new partner.

RESULTS

For both groups of students we used the difference between the percentage of the total grade that they could obtain prior to the third course component and the percentage of the total grade that they obtained posterior to the component.

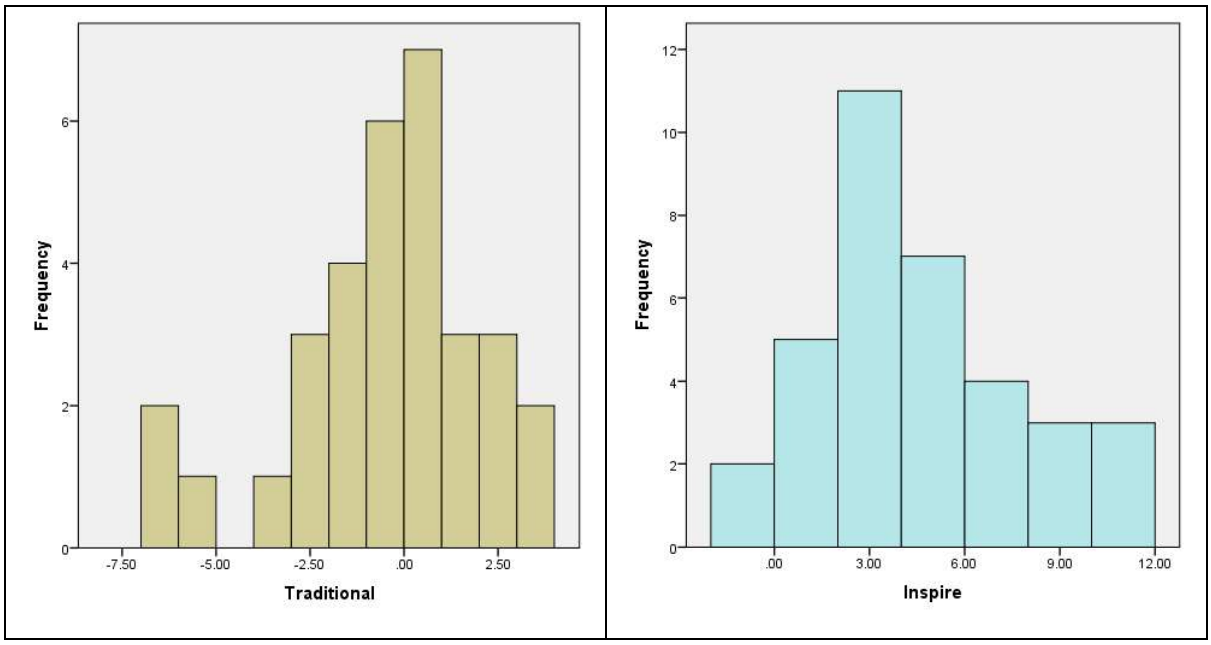


Figure 1. Histograms of relative grade difference for Traditional and Inspire groups.

There are more than 25 data points in each sample (32 for Traditional and 35 for Inspire). The distribution of grade difference (Diff) may be normal. The mean and standard deviation for the Traditional and Inspire groups are respectively 0.62 and 2.48, and 4.34 and 3.13.

Given the above and the fact that both groups may be considered as two independent samples drawn from the student population, a *t*-test for independent samples was used to verify the hypothesis. From the Levene's test for equality of variances we obtained a *p-value* = 0.094, which is greater than the typically assumed α level of 0.05. Therefore, the null hypothesis that

the variability of the two groups is equal is not rejected and the assumption of the *t*-test has been met. The result of the *t*-test indicates that the means from the two samples are significantly different ($p \leq 0.0001$); the mean difference is 4.959 with standard deviation of 0.693. This confirms our hypothesis that the replacement of the traditional Research project component with the Inspire-based component results in an average grade increase of almost 5%.

The difference between the *p*-value and the α level (0.094 vs. 0.05) may be considered small, therefore a *t*-test where variances are not assumed to be equal may be more appropriate. The *p*-value and the mean and variance values are, however, identical for equal and not equal variance values. Furthermore, running a non-parametric Mann-Whitney test for independent samples, which does not require normal distribution and large samples, also confirmed the hypothesis ($p \leq 0.000$).

DISCUSSION AND CONCLUSIONS

What could explain the difference in the learning outcomes between the two groups? One possibility is increased intrinsic motivation and engagement stimulated by the problem-solving character of negotiations and role-playing, reinforced by the fact that each student was in the centre of the learning process; each negotiated outside of class, without teacher supervision. A body of research on role-playing in negotiation courses stresses greater student engagement leading to better learning gains. If carefully planned, role-plays increase motivation and engage students "behaviourally, cognitively and emotionally" in the learning process [14]. The students' negotiation journals clearly indicate that the majority were involved in the process both emotionally and cognitively. To illustrate, here are a few quotes: "*What is he thinking?" "This is the worst offer I've received so far"*; "*After two days of waiting…I was so worried… It's pretty stressful and I felt anxious waiting for the message"*; "*I will wait until tomorrow to think about this"; "It took me about two days to reach a decision regarding the offer that Mr. Mosico sent me".* [15].

The students in the Inspire group were asked what contributed to their language improvement. They attributed it to the fact that Inspire gave them a real, authentic experience, which supports the notion of the value of experiential learning. This suggests that, asking students to do something and then analyzing the experience can be a more powerful teaching strategy than asking students to apply concepts from readings to analyzing a case.

Another contributing factor could have been the interaction with the system and its analytical decision aids over an extended period of time. According to language acquisition theory, learners need comprehensible input from texts, peers, and teachers to move to a higher level of language proficiency. Our study indicates that input from an NSS may have a powerful impact on language development.

The three components of both courses were strongly interrelated thematically and together created a rich, learning environment that enabled language acquisition. Each of the components also fostered development of critical thinking, albeit through different types of activities. However, the section that used online negotiations demonstrated better learning gains. Inspire negotiations were a new experience of practicing communication skills in an authentic setting. Joint problem solving required that the students engage in meaningful interactions, "picking up" language naturally. At the same time, however, the asynchronous aspect of the Inspire negotiations gave the learners "extended time and opportunity to react, respond, and interact with the material and each other" [16] and thus created conditions for both language acquisition and language learning making it an excellent tool for language teaching.

REFERENCES

1. M. Kersten, M. Haley and G.E. Kersten, "Developing Analytic, Cognitive and Linguistic Skills with an Electronic Negotiation System", *HICSS-36, Hawaii,* U.S.A., 2003.
2. Payne Smart, M. (December 10, 2008). *The Word and the World: Technology Aids English Language Learners.* Retrieved from https://www.edutopia.org/technology-software-english-language-learners.
3. Krashen, S. D. (1987). *Principles and Practice in Second Language Acquisition*. London, Prentice Hall.
4. Firth, A. (1995). Introduction and Overview. *The Discourse of Negotiation: Studies of Language in the Workplace*. A. Firth, Ed., New York, Pergamon Press: 3-40
5. Vygotsky, L. S. (1978). *The mind in society: The development of psychological processes*. Cambridge, MA: Harvard U. Press.
6. M. Kersten, and T. O'Brien, "A Cross-Cultural Learning Project: Learning about Cultures and Research Methods" *HICSS-44, Hawaii,* U.S.A., 2011.
7. Gwo-Jen Hwang, Hsiang Cheng et al., (2009). "Development of a Web-Based System for Diagnosing Student Learning Problems on English Tenses." In *Methods and Applications for Advancing Distance Education Technologies: International Issues and Solutions.* M.R. Syed (Ed.). IGI Global.
8. Smartsettle: http://www.smartsettle.com/home/products/smartsettle-one
9. Negoisst: http://www.wi1.uni-hohenheim.de/negoisst.html
10. Inspire, Invite Negotiation Systems, http://invite.concordia.ca
11. Burke, A. (1996). Professionalism: its Relevance for Teachers and Teacher Educators in Developing Countries. Prospects, 26 (3), 531-542, UNESCO
12. Norman, G.R. & Schmidt, H.G. (1992) The Psychological Basis of ProblemBased Learning: A Review of the Evidence. Academic Medicine, September 67 (9), 557- 565
13. Thomas, J.W. (2000). A Review of Research on Project-Based Learning. http://www.bie.org/files/researchreviewPBL_1.pdf
14. Pally, M. (2000). Sustaining interest/advancing learning: Sustained content-based instruction in ESL/EFL—Theoretical background and rationale. In M. Pally (Ed.), *Sustained content teaching in academic ESL/EFL: A practical approach* (pp. 1-18). Boston: Houghton Mifflin.
15. Poitras, J., Stimec, A. and Hill, K. (2013). On teaching. Fostering student engagement in negotiation role plays. *Negotiation J.* 29(4): 439-465.
16. Jaffee, D. (1998). Institutional Resistance to Asynchronous Learning Networks. *Asynchronous Learning Networks* 2(2): 21-31.

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

How Can I Tell You What I Want? Visualization of Principals' Preferences and Its Implications

Gregory E. Kersten
J. Molson School of Business, Concordia University
Montreal, Canada
gregory.kersten@concordia.ca

Ewa Roszkowska
Faculty of Economy and Management, University of Bialystok
Bialystok, Poland
erosz@o2.pl

Tomasz Wachowicz
Department of Operations Research, University of Economics in Katowice,
Katowice, Poland
tomasz.wachowicz@ue.katowice.pl

ABSTRACT

The problem discussed in this paper concerns principals (managers of an organization), who ask agents (their staff) to negotiate a contract with another organization. Typically, the agents negotiate with different counterparts, therefore they are likely to achieve different contracts. In addition, different perception of their principals' needs also leads to different contracts. The agents need to learn their principals' preferences which requires that the latter express these preferences in some manner. The most precise numerical preference representation may be difficult in organizational settings. Verbal and/or graphical representation is more common but susceptible to interpretation error. We used two representation formats verbal+pie-charts and verbal+bars. The results show that the both formats of preference visualization produce similar errors. Among the many individual characteristics of the agents we found that only the understanding of the negotiation case and English proficiency affect the differences between the principal's and the agent's preferences.

Keywords: Preference, Preference visualization, Preference impartation, Negotiation support, Principal-agent, Measurement

INTRODUCTION

DSS, recommender systems, and other IT-based systems that aim at aiding their users in decision-making, either require that preferences be available prior to the system's use or that they employ a preference elicitation facility that elicits preferences directly from the user. Literature offers many elicitation methods and a large number of them have been implemented in DSS and other systems. New elicitation approaches have been proposed, some based on complex computational methods, including hesitant fuzzy preference relations, interval-valued intuitionistic fuzzy numbers, and probabilistic ordered partition of super-exponential combinatorial state-space [1-3]. Less effort has been devoted to the relationship between the preference elicitation method and the quality of the decision-making process and its outcomes.

In an early study, Belton [4], compared analytic hierarchy process (AHP) and a simple multi-attribute value method (MAUM) and found that there was no strong preference for either method, though decision makers found the MAUM to be easier to understand than AHP. A recent, experimental study shows that "users consider … pairwise comparisons to be higher in decisional conflict, more effortful, less accurate, and overall less desirable to use than with absolute measurements" [5]. These and other studies (see [5, 6] for review) focus on the users' assessment of the ease of use and accuracy.

An assumption that is not often stated and which underlies many of the above-mentioned studies is that system users are homogenous. Typically, this is not the case as Fasolo and Bana e Costa's [7] experimental study shows. Given that some methods employed a numerical elicitation procedure while others employed a non-numerical procedure, the authors assessed the users' numerical literacy as well as their verbal fluency. More numerate users preferred the numerical procedure, while more verbally fluent users preferred the non-numerical procedure. Each type of users spent less time and achieved higher consistency with their preferred procedure [7].

PREFERENCES OF PRINCIPALS AND OF THEIR AGENTS

DSS, which contain a preference elicitation facility, allow their users to construct a value or utility function. In the organizational context this implies that the users' preferences are the same as the preferences of the organizations they represent. The users are the agents who act on behalf of the organizations, i.e., principals.

Managers and professionals make decisions on behalf of their organizations; they act as *agents* and the organizations as *principals*. The formulation of the most efficient contract that governs the *principal-agent relationship* is of concern to the positive agency theory. The assumption is that both actors are rational and self-interested and that the agent knows what the principal wants or needs [8]. From this assumption, it follows that the principal embeds rewards in the contract so that the agent is motivated to achieve the outcomes that the principal wants.

The assumptions of positive agency theory and its focus on extrinsic motivation have been found often inadequate for the rich and complex social-psychological profiles of the principals and their agents. To address these inadequacies, behavioral agency theory and social exchange theory have been proposed [9, 10]. The actors in behavioral agency theory are socially imbedded and their relationship is richer than in positive agency theory. In contrast with the positive agency theory, the principals may not know very well what should be done, the agents may have less information than the principals, the principals may be risk seeking, and the agents—risk-averse [11, 12]. They may have diverging interests, different views, they may misunderstand the requirements or have insufficient information.

Preference impartation

The problem discussed in this paper concerns principals, who ask agents to negotiate a contract with another organization. Typically, the agents negotiate with different counterparts, therefore they are likely to achieve different contracts. However, there are also other reasons for the agents to achieve different results. Due to the agents' different perception of principals' preferences; the principals need to express and communicate their preferences and the expression format affects the agents' comprehension. Limitations of human competence, time pressure, insufficient effort, and misunderstanding mean that aligning the agent's preferences with the principal's may not be possible [13].

In the study, multi-attribute decision problems are used because they are ubiquitous in business and government organizations. In contrast to many studies in behavioral agency theory this study focuses on two rarely researched aspects of the principal-agent problem: (1) the principals' ability to impart their preferences on the agents; and (2) the impact of the imparted preferences on the outcomes achieved by the agents. The experiments show that even in the situations in which the preferences are expressed in an unequivocal and a much more precise way than when natural language is used, the agents distort them because of their bounded rationality and genuine incompetence.

This study uses a dataset of online bilateral business negotiation experiments conducted in the Inspire system for online negotiations [14]. There are two types of principals: an entertainment agency and an artist; every agent represents one of the principals. All participants obtain information about the objectives and preferences of their principals. Then, they are asked to formulate a system of cardinal ratings (scores).

The principal's preferences were described both verbally and graphically. The verbal description was identical for all users. There were graphical representations: (1) pie charts; and (2) bars. The Inspire interface was identical for both representations and the users used the same preference elicitation facility.

Initially, the verbal representation was augmented with pie-charts. We have realized that there were significant differences between the principal's preferences and the preferences formulated by the agents [15].

We thought that the accuracy would be improved with bar graphs because data read from a bar graph, allows making judgments about differences along a common scale instead of having to make less accurate angle judgments [16]. However, the differences between the principal's and the agents' preferences also occurred with the bars-based representation.

Comparison of preference schemas

There is no single way to compare (dis-)similarity between the principal's and the agent's preferences. That is why, the two scoring systems are compared using both ordinal and cardinal similarity measures.

Because of the differences between the two negotiating sides, we used 204 data points describing the agent representing the entertainment agency. These agents were asked to determine the principal's preference based on verbal+pie representation. The ordinal Jaccard distance for pairwise comparison for these agents is 0.01, i.e., 10%. There were 32% of agents who had the same ordinal preferences their agents the average Jaccard distance for the remaining 68% agents was 0.14.

We used block distance L_1 to measure ordinal distance between the two preferential systems. The mean ordinal distance is 69.6 (σ = 48.2).

The observed differences between the communicated preferences and preferences used in the negotiations impacted the negotiation outcomes. The average difference between the agent's and the principal's rating of the agreement is 7.6% (σ = 8.2). Rating of only 4.4% of the agents was the same as their principals, while for 10% the difference was over 17.5.

There were 196 agents who were asked to determine the principal's preference based on verbal+bar representation. Despite the earlier studies that bars provide greater accuracy than pies, we found no significant difference for any of the above measurements.

Individual characteristics

There were agents who represented the principal's preference accurately and these who did not. Individual characteristics are likely to be the reason for these (in)accuracies. To determine the possible sources for the differences in elicited preferences we used two questionnaires: (1) REI Inventory used to assess intuitive–experiential and analytical–rational thinking styles [17] and (2) Thomas-Kilmann Conflict MODE Instrument used to assess negotiation styles [18]. We also used several questions about English proficiency, case understanding, prior negotiation experience, and negotiator's expectations. We found no impact of thinking styles and negotiation styles on accuracy. The results indicate that only English proficiency and case understanding have significant impact on the dissimilarity between the principal's and agents' preferences.

CONCLUSIONS

One potential difference that we did not consider is the participant's subjective graph literacy. There are differences among people in terms of their understanding of graphs. Individuals with high graph literacy have been found to make more elaborate inferences when viewing graphical displays as compared with less graph-literate individuals [19, 20]. These differences may contribute to the dissimilarity of the principal's and the agents' preferences.

REFERENCES

1. Xu, Y., et al., *Deriving the priority weights from incomplete hesitant fuzzy preference relations in group decision making.* Knowledge-Based Systems, 2016. **99**: p. 71-78.
2. Yue, C., *A geometric approach for ranking interval-valued intuitionistic fuzzy numbers with an application to group decision-making.* Computers & Industrial Engineering, 2016. **102**: p. 233-245.
3. Tran, T., D. Phung, and S. Venkatesh, *Modelling human preferences for ranking and collaborative filtering: a probabilistic ordered partition approach.* Knowledge and Information Systems, 2016. **47**(1): p. 157-188.
4. Belton, V., *A comparison of the analytic hierarchy process and a simple multi-attribute value function.* European J. of Operational Research, 1986. **26**(1): p. 7-21.
5. Aloysius, J.A., et al., *User acceptance of multi-criteria decision support systems: The impact of preference elicitation techniques.* European J. of Operational Research, 2006. **169**(1): p. 273-285.
6. Pommeranz, A., et al., *Designing interfaces for explicit preference elicitation: a user-centered investigation of preference representation and elicitation process.* User Modeling and User-Adapted Interaction, 2012. **22**(4-5): p. 357-397.

7. Fasolo, B. and C. Bana Costa, *Tailoring value elicitation to decision makers' numeracy and fluency: Expressing value judgments in numbers or words.* Omega, 2014. **44**: p. 83.
8. Weitzman, M.L., *Efficient incentive contracts.* The Quarterly J. of Economics, 1980. **94**(4): p. 719-730.
9. Bottom, W.P., et al., *Building a Pathway to Cooperation: Negotiation and Social Exchange between Principal and Agent.* ASQ, 2006. **51**(1): p. 29.
10. Westphal, J.D. and E.J. Zajac, *A behavioral theory of corporate governance: Explicating the mechanisms of socially situated and socially constituted agency.* The Academy of Management Annals, 2013. **7**(1): p. 607-661.
11. Marengo, L. and C. Pasquali, *How to get what you want when you do not know what you want: A model of incentives, organizational structure, and learning.* Organization Science, 2012. **23**(5): p. 1298-1310.
12. Pepper, A., T. Gosling, and J. Gore, *Fairness, envy, guilt and greed: Building equity considerations into agency theory.* Human Relations, 2015. **68**(8): p. 1291-1314.
13. Hendry, J., *The principal's other problems: Honest incompetence and the specification of objectives.* Academy of Management Review, 2002. **27**(1): p. 98-113.
14. Kersten, G.E. and S.J. Noronha, *WWW-based Negotiation Support: Design, Implementation, and Use.* Decision Support Systems, 1999. **25**: p. 135-154.
15. Roszkowska, E. and T. Wachowicz, *Inaccuracy in defining preferences by the electronic negotiation system users*, in *Outlooks and Insights on Group Decision and Negtiation*, B. Kaminski, G. Kersten, and T. Szapiro, Editors. 2015, Springer: Heidelberg. p. 131-143.
16. Friel, S.N., F.R. Curcio, and G.W. Bright, *Making sense of graphs: Critical factors influencing comprehension and instructional implications.* J. for Research in mathematics Education, 2001: p. 124-158.
17. Epstein, S., et al., *Individual differences in intuitive–experiential and analytical–rational thinking styles.* J. of Personality and Social Psychology, 1996. **71**(2): p. 390.
18. Thomas, K.W. and R.H. Kilmann, *Thomas-Kimann Conflict Mode Instrument*. 1974, Tuxedo: Xicom.
19. Gaissmaier, W., et al., *Numbers can be worth a thousand pictures: individual differences in understanding graphical and numerical representations of health-related information.* Health Psychology, 2012. **31**(3): p. 286.
20. Okan, Y., et al., *Individual differences in graph literacy: Overcoming denominator neglect in risk comprehension.* J. of Behavioral Decision Making, 2012. **25**(4): p. 390.

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

ICDSST 2017 on Data, Information and Knowledge Visualization in Decision Support Systems

Decision Support System for resource allocation in Brazil public universities

Martins Carolina Lino 1, 2, Zaraté Pascale 1

De Almeida Adiel Teixeira 2 Morais Danielle Costa

1 : IRIT – Toulouse Capitole University – 2 rue du Doyen Gabriel Marty 31042 Toulouse Cedex 9 – France
2 : UFPE - Production Engineering Department, Federal University of Pernambuco
Av. Acadêmico, Hélio Ramos, s/n, Cidade Universitária, Recife-PE, Brazil
carol_tcch@hotmail.com, Pascale.Zarate@irit.fr, almeida@cdsid.org.br, daniellemorais@yahoo.com.br
web-page: https:// www.irit.fr/ ; http://www.cdsid.org.br/

ABSTRACT

This study aims to present the design of a Decision Support System (DSS) for internal resource allocation in Brazil public universities, once, currently, there aren't any kind of general DSS for such a problem. To do so, the analysis is carried out by identifying the general model from the Ministry of Education and the models from every Federal University, finding similarities between each model, and, dividing the models into categories, according to their similarities. The perspectives are to contribute to the decision problem of how to allocate resources properly faced by Brazilians public universities, take safer and reliable decisions, seeking to reduce uncertainties and to maximize their results.

Keywords: Decision Support System, Design, Resource Allocation, Budgeting

INTRODUCTION

One of the ongoing challenges faced by universities in general and especially in Brazil, where public universities perform an important role, it is to improve the provision of beneficial results for the society interest, considering an increasingly complex and changing environment. Therefore, the design of a Decision Support System (DSS) for resource allocation it is an important tool to respond to this ongoing challenge.

A Decision Support System can be defined as a computer-based information system that supports decision makers use data and models to solve semi-structured and unstructured problems. It helps decision makers to make better decisions and to answer complex questions [1, 2].

Generally, considering different definitions for a DSS, they all share the idea that a DSS is essential to support the decision-making process [2] and that is the reason its applicability will be considered for this study.

Thus, this work aims to present the design of a Decision Support System (DSS) for internal resource allocation in Brazil public universities, once, currently, there aren't any general DSS for such a problem, and this can contribute to the decision question of how to allocate resources properly faced by Brazilians public universities, enabling them to take safer and reliable decisions. Also, it should be considered that public universities in Brazil use their taxpayers' money to provide education services. As a result, there is significant societal interest (or at least should exist) in the way such money is allocated, where the cost of a failure is seen as something unacceptable [3].

Within this context, it is important to clarify that the main decision of each model (not the problem situation of this study) it is how to allocate resources correctly, and the Decision Maker is considered as each Federal University.

It is known that the correct use of a DSS can improve the competences of the Decision Maker in understanding better the considered problem, how to select efficient alternatives, cost and time savings [2].

SURVEY

The design of the DSS will consist, at first, in analysing possible courses of action for the case [4]. It will involve the process of understanding the resource allocation models in public universities in Brazil, comparing them and finding similarities between the models, with the aim of generating solutions and testing feasible solutions in the future for the problem.

The general resource allocation model in Brazil is based on the "OCC Matrix" (Others, Costing and Capital Matrix). This matrix has the purpose of establishing criteria for resource allocation in Brazil's Federal Universities, and it has equitable, qualitative, inductors, measurable and auditable criteria. The model is common for all federal Universities and the structure of the budget is programmed the year before [5].

The parameters are legally defined by the Brazilian Ministry of Education (Department of Education - MEC), and the basis of the matrix is the number of students (equivalent students) from each Federal University (FU) [6]. The general model is described in Figures 3, 4 and 5.

There are 55 Federal Universities in Brazil that receive resources from the OCC Matrix, and each one of them has their own resource allocation model.

Therefore, the resource allocation process could be described by Figure 1.

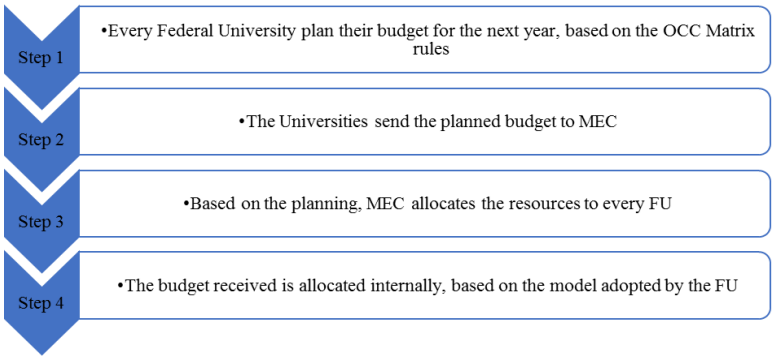


Figure 1: Resource allocation process in Brazil Federal Universities

It is important to point out that the DSS prototype proposed in this study is focused on the process described in STEP 4 (Figure 1), once the general model applicable for Steps 1 and 3 already exists (OCC Matrix). To do so, this survey was divided into three steps, that are shown in Figure 2.

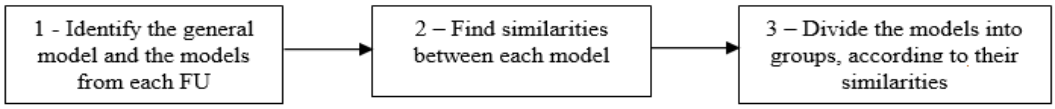


Figure 2: Steps of the survey

PART =	**h1(PTAE) + h2(EQR)**	**TAE=**	**TAEG+TAERM+TAEM+TAED**
h1=	0,9	TAEG=	∑ {[(NACG)(1+R)+N-NACG/4](PG)(DG)(BT)(BFS)}
h2=	0,1		
		TAEG = Total of Equivalent Students in Undergraduation	
PTAE =	**TAE/∑TAE**	NACG = Total of students that finished Undergraduation Studies	
PTAE = participation of the FU from the total of Equivalent Students of all the FU's		N = Total of students that starts Undergraduation Studies	
TAE = total of equivalent students		D = Duration of the undergraduation course	
		R = Standard "retention" factor of the undergraduation course	
EQR =	**DEQ/∑DEQ**	PG = weight of the undergraduation course	
EQR = efficiency and scientific academic quality from the FU		BT = bonus for having nightly undergraduation courses	
DEQ = efficiency and scientific academic quality dimension from the FU		BFS = bonus for having an undergraduation course outside the main campus	
∑DEQ = efficiency and scientific academic quality dimension from the set of FU's			
		TAEG**= ∑(NMG)(PG)(BT)(BFS)	
DEQ = DEAE + DQG + DQM + DQD		**= new undergraduation courses (less than 10 years)	
DEAE = efficiency dimension of the teaching activities in the FU		NMG = Total of students enrolled in an undergraduation course	
DQG = quality dimension from the undergraduation courses		PG = weight of the undergraduation course	
DQM = quality dimension from the master courses		BT = bonus for having nightly undergraduation courses	
DQD = quality dimension from the doctorate courses		BFS = bonus for having an undergraduation course outside the main campus	

Figure 3: General model – part 1

DEAE = FRAP		TAEG***= ∑ {[(NACG)(1+R)](PG)(DG)(BT)(BFS)}
		***= New undergraduation course
FRAP = RAP / RAP		
FRAP = relation factor between equivalent student and professor		DG = Standard duration of the undergraduation course
RAP = relation between equivalent student and professor		
RAP = average relation between equivalent student and professor		
		DQM = ∑FQM / NCM
DQG = ∑FCG / NCG		FQM = (CCM / CCM)
		FQM = quality factor from the master course
FCG = (CSG / CSG)		NCM = total number of master courses at the FU
FCG = quality factor from the undergraduation course		CCM = CAPES concept of the master course
CSG = SINAES concept of the undergraduation course		CCM = average CAPES concept from the set of FU's of the master courses
CSG = SINAES average concept from the undegraduation course from the set of FU's		that have the same area
NCG = number of undergraduation courses evaluated at the FU		
		DQD = ∑FQD / NCD
		FQD = (CCD / CCD)
		FQD = quality factor from the doctorate course
		NCD = total number of doctorate courses at the FU
		CCD = CAPES concept of the doctorare course
		CCD = average CAPES concept from the set of FU's of the doctorate
		courses that have the same area

Figure 4: General model – part 2

TAERM = ∑(NAMRM)(PRM)	TAED = ∑(NACD)(DD)(PD)
TAERM = total of equivalent students from medical residency	TAED = total of equivalent students in a doctorate course
NAMRM = total of students enrolled in a medical residency course	NACD = total of students that concluded the doctorate course
PRM = weight of the group from the medical residency course	DD = standard duration of the doctorate course
	PD = weight of the group from the doctorate course
TAEM = ∑(NACM)(DM)(PM)	
TAEM = total of equivalent students in a master course	
NACM = total of students that concluded the master course	
DM = standard duration of the master course	
PM = weight of the group from the master course	

Figure 5: General model – part 3

When analysing the available models (only 30 models were available for consulting or the university doesn't have a defined model) and their similarities, it was possible to divide them into three main categories: Model 1, based on the general resource allocation model, Model 2, based on some indicators suggested by the Brazilian audit office (Tribunal de Contas da União - TCU) [7], and, Model 3, based on some indicators that will be shown next.

MODEL 1

Model 1 is based on the general resource allocation model presented in Figures 3, 4 and 5, but some universities vary or adapted a few parameters from it.

MODELS 2 and 3

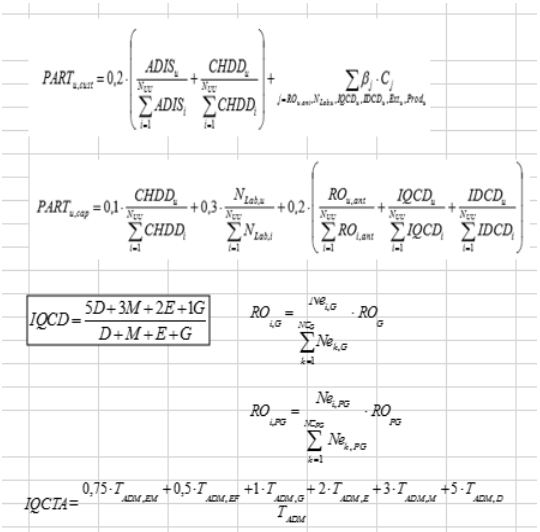


Figure 6: Indicators from model 2

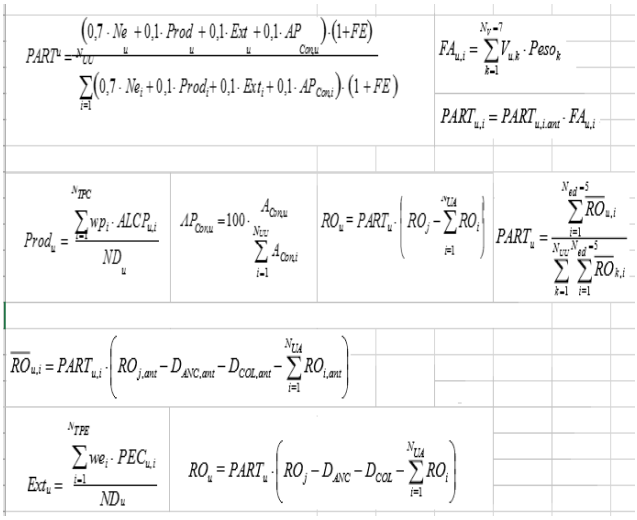


Figure 7: Indicators from model 3

Model 2 is based mainly on indicators like costing; the amount of hour of each course; the number of students in every course; the number of professors and their workloads in teaching, research and extension activities; publications from every academic department; the number of laboratories and qualification of the academic staff.

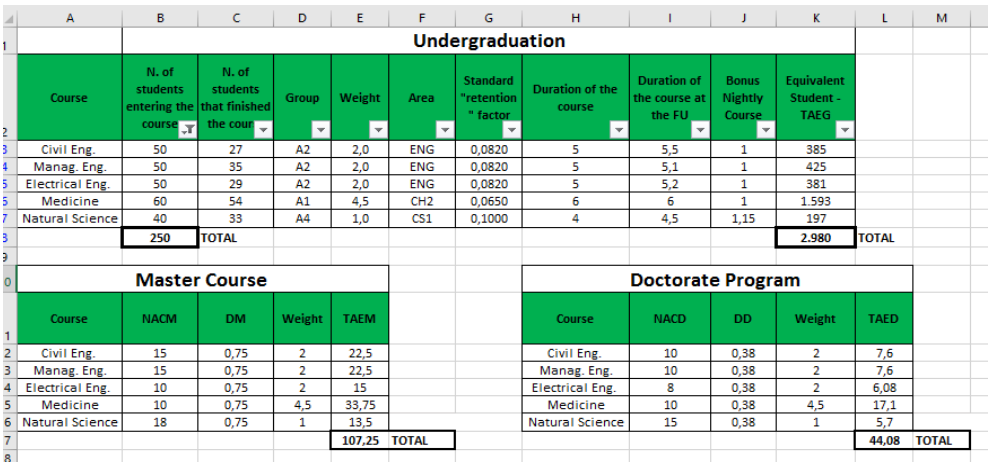
Model 3 is based basically on the following indicators: number of professors; the number of technical employees; the number of students from each department; the total area from the laboratories; the total area from the departments; scientific production from the

departments; extension activities and others.

DSS PROTOTYPE

The DSS Prototype from the main three models found by this study is presented next. The models were divided into categories, according to their similarities. This initial prototype was designed in a Microsoft Excel file and it was the first step of a bigger research, that aims to improve the design of this DSS, by transforming the prototype into a web-based system, with a programming language, developing the data basis for the model and for the users, implementing the program, and, finally, tested by the users. The research also will include a project portfolio selection approach as an appropriate model to analyze the resource allocation process of the universities.

Model 1



Undergraduation

Course	N. of students entering the course	N. of students that finished the cour	Group	Weight	Area	Standard "retention " factor	Duration of the course	Duration of the course at the FU	Bonus Nightly Course	Equivalent Student - TAEG
Civil Eng.	50	27	A2	2,0	ENG	0,0820	5	5,5	1	385
Manag. Eng.	50	35	A2	2,0	ENG	0,0820	5	5,1	1	425
Electrical Eng.	50	29	A2	2,0	ENG	0,0820	5	5,2	1	381
Medicine	60	54	A1	4,5	CH2	0,0650	6	6	1	1.593
Natural Science	40	33	A4	1,0	CS1	0,1000	4	4,5	1,15	197
	250	TOTAL								2.980 TOTAL

Master Course

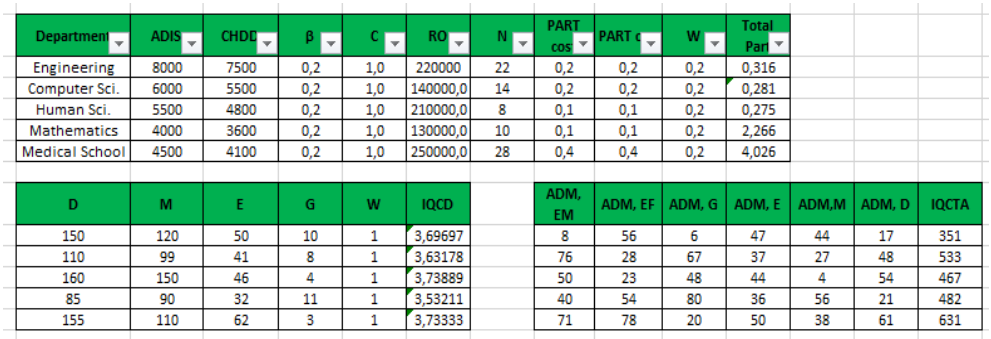
Course	NACM	DM	Weight	TAEM
Civil Eng.	15	0,75	2	22,5
Manag. Eng.	15	0,75	2	22,5
Electrical Eng.	10	0,75	2	15
Medicine	10	0,75	4,5	33,75
Natural Science	18	0,75	1	13,5
				107,25 TOTAL

Doctorate Program

Course	NACD	DD	Weight	TAED
Civil Eng.	10	0,38	2	7,6
Manag. Eng.	10	0,38	2	7,6
Electrical Eng.	8	0,38	2	6,08
Medicine	10	0,38	4,5	17,1
Natural Science	15	0,38	1	5,7
				44,08 TOTAL

Figure 8: DSS Prototype – Model 1

Model 2

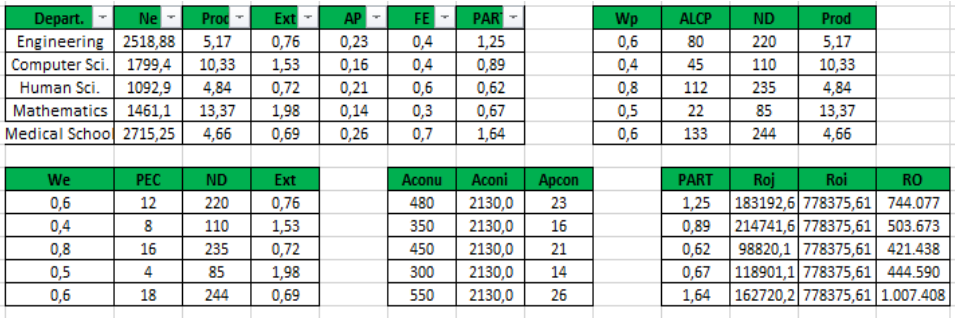


Department	ADIS	CHDD	β	C	RO	N	PART cos	PART c	W	Total Par
Engineering	8000	7500	0,2	1,0	220000	22	0,2	0,2	0,2	0,316
Computer Sci.	6000	5500	0,2	1,0	140000,0	14	0,2	0,2	0,2	0,281
Human Sci.	5500	4800	0,2	1,0	210000,0	8	0,1	0,1	0,2	0,275
Mathematics	4000	3600	0,2	1,0	130000,0	10	0,1	0,1	0,2	2,266
Medical School	4500	4100	0,2	1,0	250000,0	28	0,4	0,4	0,2	4,026

D	M	E	G	W	IQCD	ADM, EM	ADM, EF	ADM, G	ADM, E	ADM,M	ADM, D	IQCTA
150	120	50	10	1	3,69697	8	56	6	47	44	17	351
110	99	41	8	1	3,63178	76	28	67	37	27	48	533
160	150	46	4	1	3,73889	50	23	48	44	4	54	467
85	90	32	11	1	3,53211	40	54	80	36	56	21	482
155	110	62	3	1	3,73333	71	78	20	50	38	61	631

Figure 9: DSS Prototype – Model 2

Model 3



Depart.	Ne	Proc	Ext	AP	FE	PAR	Wp	ALCP	ND	Prod
Engineering	2518,88	5,17	0,76	0,23	0,4	1,25	0,6	80	220	5,17
Computer Sci.	1799,4	10,33	1,53	0,16	0,4	0,89	0,4	45	110	10,33
Human Sci.	1092,9	4,84	0,72	0,21	0,6	0,62	0,8	112	235	4,84
Mathematics	1461,1	13,37	1,98	0,14	0,3	0,67	0,5	22	85	13,37
Medical Schoo	2715,25	4,66	0,69	0,26	0,7	1,64	0,6	133	244	4,66

We	PEC	ND	Ext	Aconu	Aconi	Apcon	PART	Roj	Roi	RO
0,6	12	220	0,76	480	2130,0	23	1,25	183192,6	778375,61	744.077
0,4	8	110	1,53	350	2130,0	16	0,89	214741,6	778375,61	503.673
0,8	16	235	0,72	450	2130,0	21	0,62	98820,1	778375,61	421.438
0,5	4	85	1,98	300	2130,0	14	0,67	118901,1	778375,61	444.590
0,6	18	244	0,69	550	2130,0	26	1,64	162720,2	778375,61	1.007.408

Figure 10: DSS Prototype – Model 3

PERSPECTIVES AND CONCLUSIONS

The purpose of this study was to present the design of a Decision Support System (DSS) for internal resource allocation in Brazil public universities. To do so, the survey was divided into three steps: identify the general model and the models from each FU; find similarities between each model; and, divide the models into categories, according to their similarities. This initial prototype was the first step of a bigger experiment. The system still must be improved to be useful for the users.

The next step is to transform the DSS prototype into a web-based system, with a programming language, constructing its data basis for the model and for the users, implement the program, and, finally, tested by the users. Also, the DSS could have potential expansions in the future, expanding its general prototype to be used by the Ministry of Education in Brazil or others public institutions with the similar decision problem.

The perspectives are to contribute to the decision problem of how to allocate resources correctly faced by Brazilians public universities, take safer and reliable decisions, seeking to reduce uncertainties and to maximize their results. In addition, it could be used to provide background for the Federal Universities strategic resource allocation planning.

It is worthwhile to note that the DSS prototype has no production intention but to deal with as an experiment with only research purposes.

REFERENCES

1. Bidgoli, H. “Decision Support System - Principles and Practice.” New York: West Publishing Company, 1989.

2. Sprague, Jr.; Watson, H. “Decision support systems: putting theory into practice.” USA: Prentice-Hall, 1989.

3. Williams, G. “Applying Management of Risk (M_o_R®) for Public Services.” GSW Consultancy, 2009.

4. Zaraté, P. “The process of designing a DSS: A case study in planning management.” Eur. J. Oper. Res. 55, 394 – 402 (1991).

5. Brasil. “Decreto nº 7.233 de 19 de julho de 2010”. Procedimentos orçamentários e financeiros relacionados à autonomia universitária, e dá outras providências. Brasília. Available in: http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2010/Decreto/D7233.htm

6. Ministério da Educação (MEC). “Portaria nº 651 de 24 de julho de 2013”. Institucionaliza, no âmbito do Ministério da Educação a Matriz de Orçamento de Outros Custeios e Capital- Matriz OCC. Brasília. Available in: http://sintse.tse.jus.br/documentos/2013/Jul/26/portaria-no-651-de-24-de-julho-de-2013-fica

Tribunal de Contas da União (TCU). “Orientações para o Cálculo dos Indicadores de Gestão. Decisão TCU nº 408-2012 – Plenário”. Brasília. 2012

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

A Loose-Coupling Integration of the MACBETH Approach in ArcGIS

Francis Marleau Donais, Roxane Lavoie
Graduate school of Land Management and Urban Planning, Université Laval
2325 allée des Bibliothèques , Québec G1V 0A6, Canada
francis.marleau-donais.1@ulaval.ca, roxane.lavoie@esad.ulaval.ca

Irène Abi-Zeid
Department of Operations and Decision Systems, Université Laval
2325 rue de la Terrasse, Québec G1V 0A6, Canada
Irene.abi-zeid@osd.ulaval.ca

ABSTRACT

Spatial decision-making involving georeferenced data and multiple conflicting objectives often relies on approaches that combine multicriteria decision analysis (MCDA) methods with geographic information systems (GIS). Although some tools and plugins integrating MCDA methods such as AHP or ELECTRE in a GIS software are available, none of these have been developed for MACBETH, another popular MCDA method. The objective of this project was to develop a first set of tools to convert results obtained using MACBETH into vector data presented as layers in ArcGIS. Using the ArcGIS visual programming language and Python, we built three types of tools: spatial aggregation tools, a criterion standardization tool, and a multicriteria aggregation tool. In this paper, we present these tools and illustrate their applicability in an urban transportation planning project. The project consisted of the multicriteria evaluation of streets' potential to be redesigned as complete streets in Quebec City, Canada. A total of 5019 street segments were assessed according to 11 criteria representing environmental, transportation, and urban planning perspectives. The application showed that the developed tools were fast, easy to use, flexible and transparent. Our future research aims at fully integrating MACBETH in a GIS and developing a user-friendly spatial multicriteria decision analysis software.

Keywords: GIS-MCDA integration, MACBETH, spatial decision support systems, transportation planning

1. INTRODUCTION

Geographic information systems (GIS) are decision support systems (DSS) that incorporate spatial information to support decisions based on georeferenced data. In general, GIS are not designed to properly aggregate multiple criteria evaluations in the context of conflicting preferences and objectives [1]. To overcome these weaknesses, a GIS can be combined with a multicriteria decision analysis (MCDA) method. This field of study, known as GIS-MCDA or multi-criteria spatial decision support systems (MC-SDSS), has generated a growing interest over the past few years [2, 3].

GIS-MCDA are tools that transform and combine geographic data with decision makers' (DM) preferences resulting in an output that can be visualized as a decision map [2]. In order to characterize the various GIS-MCDA tools, Malczewski [4] has proposed two classifications one of which is a spatial classification based on the nature of the data models (vector data or raster data), as well as on the spatial nature of the decision alternatives and the evaluation criteria (spatially implicit or explicit). Furthermore, he classified the GIS-MCDA integrations along two dimensions that he called the extent and the direction [4]. The extent of integration can be loose-coupling (two systems that exchange data as output-input), tight-coupling (two systems that exchange data and share a common user-interface) or full integration. As for the direction of integration, this can be a one-directional integration where either the GIS or the MCDA method is the main software, a bi-directional integration where the information travels as a one-time flow between a GIS and a MCDA, or a dynamic integration where the information continuously flows in both directions according to the user's needs.

Despite the growing literature on MCDA in spatial decision-making, only a few tools and plugins integrating MCDA in a GIS are currently available. These include tools developed for QGIS [5–8] or for ArcGIS [9]. Furthermore, none of these tools were developed for MACBETH, a MCDA method that aggregates an alternative's evaluations on various criteria in a weighted average score, using pairwise comparisons to build interval scales and scaling constants [10]. Since MACBETH has recently been applied in several GIS-MCDA case studies [11–14], there is a growing need to develop tools to facilitate the use of MACBETH in conjunction with a GIS, both for researchers and for practitioners, who are often novice users of MCDA methods. In this paper, we describe the first steps towards fulfilling this need.

In order to test the developed tools, we applied them to a multi-criteria spatial project using sustainable transportation principles for street rehabilitation in Quebec City, Canada. Our objective was to evaluate and rank streets that could be redesigned as complete streets, and to present their complete street potential in a GIS. Complete streets are streets designed to be accessible, safe and comfortable for everyone regardless of their transportation mode or physical condition [15]. Therefore, the development of a MC-SDSS was essential to facilitate and propose a systematic multicriteria evaluation procedure for all Quebec City streets.

This short paper is organized as follows: Section 2 presents the MACBETH method and its integration in an ArcGIS environment. Section 3 describes the case study in which the developed tools were tested. Section 4 explains the application of the tools. Finally, Section 5 discusses the tools' strengths and weaknesses as well as future research to enhance them.

2. METHODOLOGY

The process of building a MC-SDSS can be described as a five-step process [16]: (1) designing the decision process, (2) structuring the MC-SDSS, (3) eliciting spatial standardization preference functions, (4) aggregating the local preference values of the

alternatives (attractiveness of an alternative's performance on a given criterion) using criteria scaling constants (qualitative swing weights) [17], and (5) analyzing the results and making recommendations . In this project, we focused on building preference standardization functions (interval scale value functions that translate raw criteria performances into a common unit called attractiveness) and on the aggregation of the local attractiveness values of the alternatives using the MACBETH method. We chose MACBETH because it can easily rank a large set of alternatives, it helps build interval scales, and it is a user-friendly method supported by software, and is based on linguistic evaluations and weighted averages. ArcGIS was used to display the spatial aggregated results obtained with MACBETH.

In MACBETH, functions that are built using the DM preferences, standardize the various criteria performances into a common unit scale, called the attractiveness scale. To ensure that attractiveness scales are interval scales, the concept of *good* and *neutral* reference levels is introduced for each criterion. The *good* reference represents a performance echelon with which a DM is satisfied whereas the *neutral* reference represents a performance echelon that the DM wishes to achieve as a minimum. The difference in attractiveness between two echelons is then obtained from the DM according to a 7 point semantic scales: null, very weak, weak, moderate, strong, very strong and extreme. On these new scales, the *good* reference is assigned a value of 100 and the *neutral* reference a value of 0. The local attractiveness functions defined for each criterion are then aggregated in a weighted average using scaling constants to convert local value scores into an overall score. This approach is similar to the swing-weighting technique of SMART, however it uses qualitative information rather than quantitative information. The scaling constants are elicited following pairwise comparisons of fictitious alternatives and using the previous semantic scales. The DM is asked to express his/her preferences and strengths of preference about the swing from the *neutral* level to the *good* level of criterion *i* compared to the swing from the *neutral* level to the *good* level of criterion *j*. All the DM judgements are computed in the M-MACBETH software to infer the attractiveness scales and the scaling constants. Although very helpful in supporting a multicriteria decision analysis process, this software is however, in its current form, not well suited to handle spatial data.

The spatial nature of the data in our transportation project required the use of a GIS to assess the alternatives. As a first attempt to combine MACBETH in a GIS, we developed a set of tools with a loose-coupling integration and a one directional software approach, with ArcGIS as the main software. The developed tools use the output from M-MACBETH as input to ArcGIS. They were developed in ArcGIS because the Quebec City planners involved in our case study were familiar with ArcGIS.

In order to rank the alternatives (streets) as a function of their potential to be redesigned as complete streets, we developed three tools: a spatial aggregation tool (SAT), a criterion standardization tool (CST), and a multicriteria aggregation tool (MAT). To illustrate how these tools work, we will use the bus network criterion as an example. First, for this criterion, SAT overlaid bus network layers on a map containing every alternative, namely the Quebec City streets. For qualitative criteria, SAT converted qualitative scales into ordinal numerical scales (e. g., the bus network criterion has ten echelons; therefore, echelons were attributed a number from zero to nine). Second, CST was used to translate the bus network ordinal values into attractiveness values using the attractiveness functions built with M-MACBETH. Finally, MAT weighted the bus network criterion's attractiveness values according to its scaling constant and added it to the other weighted criteria's attractiveness values to produce a synthesized decision map.

To take into account data heterogeneity, each criterion required the development of a specific SAT, developed within ModelBuilder, the ArcGIS visual programming language. As for the CST and MAT tools, they were written in the Python language and were designed to handle all the criteria.

All the tools have an interface similar to other ArcGIS tools and work with the definition of input and output features (shapefiles or geodatabases). Depending on the tool, configurable parameters are included. To aid new users, a help section was incorporated to explain the general purpose of each tool and to precisely describe each parameter.

3. CASE STUDY

In most cities, street rehabilitation is an engineering-centered approach based on infrastructure obsolescence that does not take into account multidimensional aspects. However, the recent trend towards sustainable transportation planning has led Quebec City planners to question their current practices. They expressed the wish to include environmental, transportation and urban planning aspects in the decision-making process that leads to the choice of a street to be redesigned as a complete street. From a more practical perspective, they wished to have a map synthesizing the different experts' points of view that will allow them to easily identify the streets with the highest complete street potential in each neighborhood [18].

To help them achieve their goal, we designed and applied a process similar to the five-step process for building a MC-SDSS proposed by [17]. A total of six group workshops and of eleven subgroup workshops were held to structure, elicit and analyze the DMs preferences. The DMs consisted of 11 Quebec City professionals from various departments including environment, urban planning, engineering and transportation. During these workshops, we used MACBETH to help the professionals identify 11 criteria with their measurement units, build attractiveness scales for each criterion, and elicit scaling constants for the criteria. The first two workshops allowed us to structure the problem, the following three workshops aimed at applying MACBETH, and the last workshop enabled us to analyze and validate the model results. During these workshops, the authors acted as facilitators and used decision conferencing to ensure a smooth process [19]. The decision alternatives were discrete and spatially explicit since they consisted of street segments, namely portions of streets between two adjacent intersections. As for the spatial nature of the evaluation criteria, it was a mix of explicit (e.g. bicycle paths, street right of way) and implicit criteria (spatial data that needed to be processed, e.g. connectivity, density). The datasets were provided by different Quebec City departments and were heterogeneous. The only common feature between the dataset was the spatial location. All the spatial data were vector data: they were polylines directly representing a street segment, polylines parallel to a street segment or polygons that included several street segments.

4. APPLICATION

We describe briefly in what follows the application of the SAT, CST, and MAT tools to integrate the results of M-MACBETH in ArcGIS. First, using the SAT, 11 criteria maps were created. The criteria maps contained all Quebec City street segments, their unique identification (ID) number (primary key), and the non-standardized criterion values. Input and output data files were defined by the users. CST was next used to standardize the criteria maps by transforming them into 11 attractiveness maps containing the attractiveness value of

each alternative (street). CST's interface consists of a data section and a MACBETH standardization section (Figure 1). The data section includes the input and output data files, and allows the user to choose the dataset field to be standardized. The configuration must include an input row for the non-MACBETH values and an output row for the attractiveness values. The information in the rows is taken from M-MACBETH and must be typed in the right order to standardize the data appropriately.

Finally, MAT aggregated the 11 attractiveness maps into a complete street decision map containing street ID, street name, standardized value and the aggregated value. The attractiveness maps must be added one by one in the MAT configuration table and their associated scaling constants need to be entered in the weight row. The scaling constant values (weights) must be between zero and one; if the constant sum is not equal to one, an error message will appear. MAT's interface is similar to CST's interface with input and output data files; the main difference is in the configuration table.

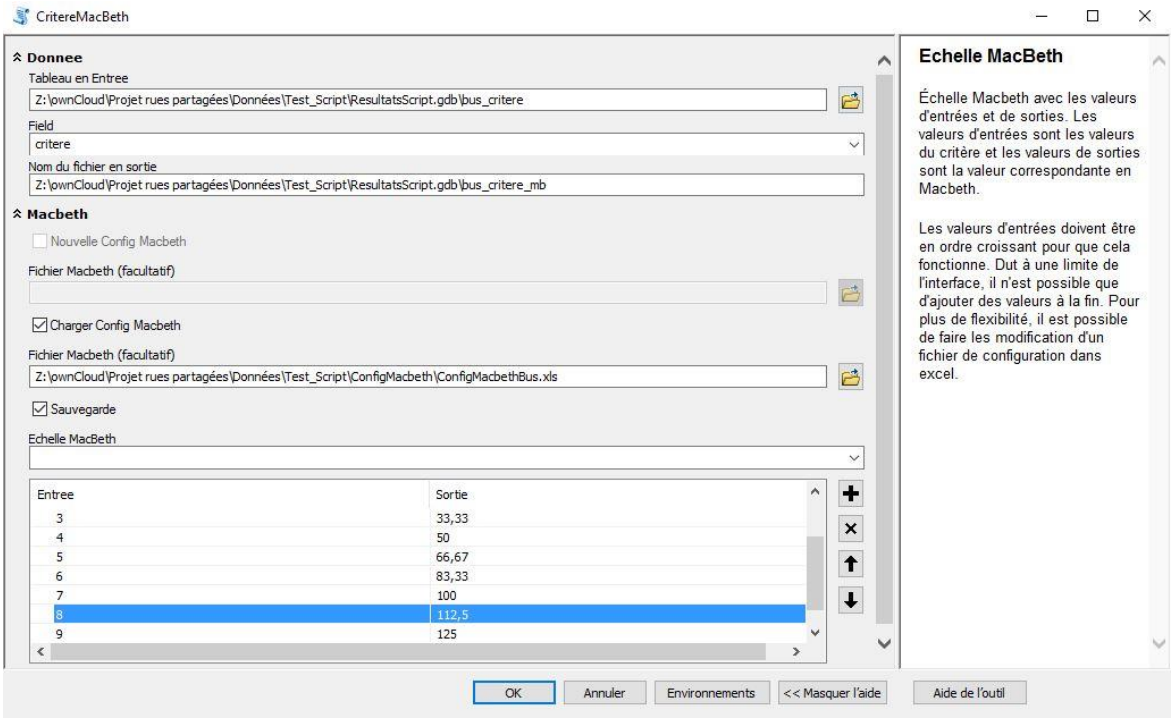


Figure 1: Criterion standardization tool in ArcGIS

As a result of applying our methodology and tools, 5019 street segments were ranked in ArcGIS based on 11 criteria. To take into account the spatial variability between the various Quebec City neighborhoods, results were computed on a neighborhood scale level. Figure 2 shows an example of the complete street potential for the neighborhood of Saint-Roch.

5. DISCUSSION AND CONCLUSION

This paper presented ongoing work to integrate the MCDA method MACBETH in ArcGIS. To achieve this, we developed a set of tools to rank spatially defined alternatives and to produce georeferenced decision maps. The developed tools are meant to be intuitive for new users since their interfaces are similar to inbuilt ArcGIS interfaces. When compared to results obtained with manual aggregation, the creation of these tools allowed us to decrease the number of manipulation errors and to speed up the computing process. As an example, the

final aggregation script took around 45 seconds to aggregate 5019 alternatives for 11 criteria compared to approximately two hours when done manually.

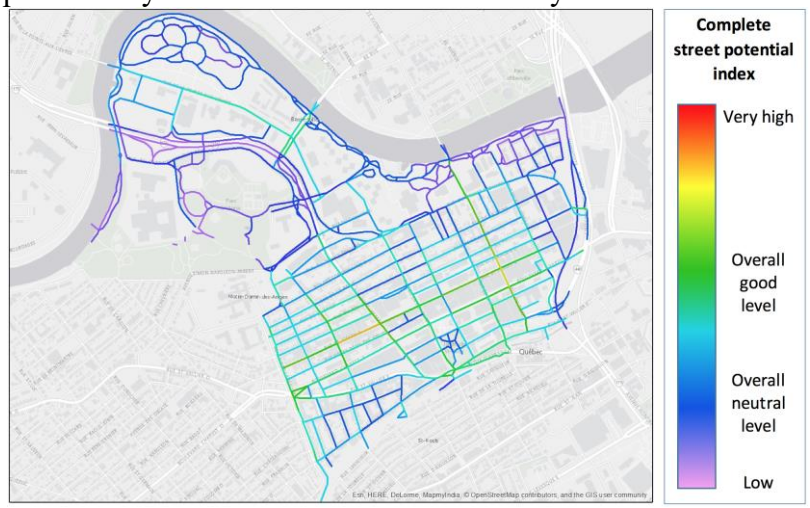


Figure 2: Complete Street potential for Saint-Roch neighborhood

Even though the CST and MAT were initially developed for integrating MACBETH in ArcGIS, they are easily adaptable and may be applied to other contexts. For example, MAT could be used to compute weighted sums in ArcGIS. However, this is not the case for the SATs, specifically developed for our case study and not easily transposable to other projects. The division of our results production procedure in several tools improves flexibility and reduces the black box effect by informing users of every step that leads to a synthesized score. However, this division increases the number of manipulations and novice users might be overwhelmed by the different tools.

One limitation of the tools is that they only process vector data and cannot handle raster data. In addition, they were developed in French because the project was in collaboration with a French-speaking city; this may create language barriers with foreign-language speakers.

Future research and development could address these issues and develop new functionalities. For example, the addition of a tool to visualize the alternatives' performances on the various criteria using bar diagrams could improve the decision-making process by providing disaggregated information to the DM. This information is already available through the attribute tables, but is not user-friendly. The coupling of the M-MACBETH software and ArcGIS should also be improved. At the current time, M-MACBETH does not have any export functions to other formats. Therefore, all M-MACBETH results had to be typed in, one by one, in the configuration tables, which is time consuming. Creating an M-MACBETH export file containing the standardization and criteria weight configurations could speed up the integration in GIS. Our ultimate goal is to develop a plugin native to ArcGIS or QGIS in order to fully integrate the MACBETH approach in a GIS.

The results of this project were very well received by the Quebec City professional who participated in the project as well as by elected officials. It has also generated a lot of interest among Quebec City citizens and received a good general press coverage in local newspapers and local television news.

ACKNOWLEDGMENTS

The authors wish to thank all the participants from Ville de Québec and UMR Sciences Urbaines for their input and valuable time. The authors are also grateful for the financial

support provided for this project by Thales, NSERC, MITACS, and Ville de Québec. Furthermore, the precious collaboration of Jean-Philippe Delisle is acknowledged.

REFERENCES

1. F. Joerin, M. Thériault, A. Musy, "Using GIS and outranking multicriteria analysis for land-use suitability assessment", Int. J. Geogr. Inf. Sci. **15**(2), 153–174, 2001.
2. J. Malczewski, "GIS-based multicriteria decision analysis: a survey of the literature", Int. J. Geogr. Inf. Sci. **20**(7), 703–726, 2006.
3. V. Ferretti: "Integrating Multicriteria Analysis and Geographic Information Systems: a survey and classification of the literature". Presented at the 74th Meeting of the European Working Group "Multiple Criteria Decision Aiding", Yverdon, Switzerland, 2011.
4. J. Malczewski, C. Rinner, "Multicriteria Decision Analysis in Geographic Information Science". Springer Berlin Heidelberg, Berlin, Heidelberg, 2015.
5. O. Sobrie, M. Pirlot, F. Joerin, "Intégration de la méthode d'aide à la décision ELECTRE TRI dans un système d'information géographique open source". Rev. Int. Géomat. **23** (1), 13–38, 2013.
6. G. Massei, L. Rocchi, L. Paolotti, A. Boggia, "Development of multicriteria modules for environmental valuation in GRASS GIS". Aestimum. **0**(63), 129–143, 2013.
7. G. Massei, "QGIS Plugin: VectorMCDA", 2014.
8. M. Selim Bilgin, "QGIS Plugin: Easy AHP", 2015.
9. C. Rinner, S. Voss, "MCDA4ArcMap–An Open-Source Multi-Criteria Decision Analysis and Geovisualization Tool for ArcGIS 10". Feature Artic. Cartouche Newsl. Can. Cartogr. Assoc. **86**, 12–13, 2013.
10. C.A. Bana E Costa, J.M. De Corte, J.-C. Vansnick, "MACBETH". Int. J. Inf. Technol. Decis. Mak. **11**, 359–387, 2012.
11. C.A. Bana E Costa, P.A. da Silva, F.N. Correia, "Multicriteria evaluation of flood control measures: The case of Ribeira do Livramento". Water Resour. Manag. **18**(3), 263–283, 2004.
12. O.E. Demesouka, A.P. Vavatsikos, K.P. Anagnostopoulos, "Using MACBETH Multicriteria Technique for GIS-Based Landfill Suitability Analysis". J. Environ. Eng. **142**(10), 1–11, 2016.
13. R. Lavoie, J. Deslandes, F. Proulx, "Assessing the ecological value of wetlands using the MACBETH approach in Quebec City. J. Nat. Conserv. **30**, 67–75, 2016.
14. R. Lavoie, F. Joerin, J.-C. Vansnick, M.J. Rodriguez, "Integrating groundwater into land planning: A risk assessment methodology". J. Environ. Manage. **154**, 358–371, 2015.
15. B. McCann: "Completing our streets: the transition to safe and inclusive transportation networks. Island Press, Washington, United States, 2013.
16. V. Ferretti, G. Montibeller, "Key challenges and meta-choices in designing and applying multi-criteria spatial decision support systems". Decis. Support Syst. **84**, 41–52, 2016.
17. C.A. Bana e Costa, J.C. Lourenço, M.P. Chagas, J.C. Bana e Costa, Development of Reusable Bid Evaluation Models for the Portuguese Electric Transmission Company. Decis. Anal. **5**(1), 22–42, 2008.
18. F. Marleau Donais, R. Lavoie, I. Abi-Zeid, J.-P. Delisle, "Évaluation du potentiel des rues à être aménagées en rues conviviales - Une approche par analyse multicritère". Université Laval, Quebec City, Canada, 2016.
19. L.D. Phillips, "Decision Conferencing". In: W. Edwards, R.F.J. Miles, and D. von Winterfeldt, (eds.) Advances in Decision Analysis. pp. 375–399. Cambridge University Press, Cambridge, United Kingdom, 2007.

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

Spatial Data Visualization on Disaggregated Property Felonies

Thyago C. C. Nepomuceno, Ana Paula Cabral C. Seixas
Universidade Federal de Pernambuco
Av. Prof. Moraes Rego, 1235. CEP: 50670-901, Recife - PE, Brazil
nepomuceno@dis.uniroma1.it, apcabral@ufpe.br
web-page: http://www.gpsid.org.br/

ABSTRACT

This work aims to meet the gap of aggregating spatial data of different criminal incidences in order to provide more accurate information for public safety strategies. The criminal patterns of property felonies are identified in a web big data curation process with the support of a Geoinformation System (GIS) to perform optimized hot-spots analysis. The felonies are typified (disaggregated) into theft, robbery, express kidnapping, burglary, saucy bank, motor vehicle robbery, steaming and attempted robbery, and the city of Recife, one of the most important Brazilian capitals, is selected for the analysis. Our results provide an integrated spatial visualization of felonies associated with environmental factors to support more efficient allocation of public safety resources.

Keywords: Criminal Behavior, Spatial Analysis, Geoinformation System (GIS), Data Visualization, Optimized Hot-spots

INTRODUCTION

A substantial number of researchers have addressed to Geoinformation Systems (GIS) to find a solid foundation for innovative strategies and public policies. These sort of technology is able to provide a valuable information that combines social, economic, contextual, demographic and spatial characteristics into an integrated tool that has been used to mitigate environmental risk, identify market niches, predict urban growth, disease outbreaking and to frame industrial clusters and social aggregations to support decision making by policy managers. The spatial incidence of crime over urban areas is another embracing subject that has been exhaustively put under discussion in recent technological advances [1, 2, 3]. The hot spot analysis on the incidence of crime with the support of Geoiformation Systems has been traditionally employed to identify spatial patterns of the criminal behavior that is not either evenly neither randomly distributed along the space [4, 5, 6 and 7]. The spatial concentration of felonies and misdemeanors receives special attention for its property to provide statistical measures that support public policy makers to design strategies based on an empirical understanding of the environmental relations that undergoes in their urban space, and

valuable information can be extracted from crime associations with exogenous determinants of criminality among different groups of individuals.

Most of traditional and recent studies, however, frame such clustering behavior with feature placing based on an aggregated dataset of property or violent crimes, without a meticulous discrimination of which sort of felony the results are limited to. Streets robberies committed in crowded places to a random passerby might present different motivations and association with different spatial characteristics than street robberies of which victims are approached closer to their residences. Whereas the first is strictly related to the opportunity granted by the victim's careless behavior or by a set of favorable environmental factors observed in the eyes of the perpetrator, the last might be more consistently associated to a premeditated property crime, less related to the opportunity and more related to strategies adopted by the criminal and prior knowledge of the victim's routine and response. Despite the importance of felonies disaggregation, this empirical assessment is rare to find since it requires a high level of detailing hardly provided by police datasets, statics or reports.

In this work such spatial evaluation is obtained resorting to an online big dataset of property offenses recorded by the victims. The concept of property crime data is decomposed into small sets of 10 felonies and misdemeanors and four instances of analysis is presented in this work. The main goal with this methodology is to provide substantial information that supports the decision making of public safety authorities by i. a rapid visualization of hot spots that indicates not only a higher incidence, but a criminal pattern that might advocate an efficient allocation of sworn officers, and ii. correlating this information of spatial concentration with environmental determinants in order to measure the impact of each potential factor in the incidence of crime. For the first consideration, the optimized hot spot analysis is employed to visualize the spatial clusters in the desegregated data that present stronger associations with the environment than larger datasets. In regards to the potential environmental determinants, some of the exogenous factors to be considered are the neighborhood density, public illumination, paving, public spaces (squares, lakes, recreation centers), urban space, income and the proportion of rented houses. Nevertheless, due the length restriction of this proposal, this later assessment is taken for granted.

It is expected a higher sensibility (response) from detailed felonies with regard the spatial association than aggregate data of property crimes, which must contribute to lead different useful strategies' development and the creation of specific approaches for specific misdemeanors. In addition, some of the analyzed crimes are very specific in Brazilian social context, such as *sequestro relânpago* (lighting kidnap), *arrastão* (flashrobs) and *saidinha de banco* (saucy bank), which makes this assessment unique due the difficult to find considerable incidence of those felonies in any country other than Brazil, thus, becoming impossible any statistical inference. The next sections describe the city of analysis, the data decomposition, methodology applied and the spatial assessment. Lastly, some discussion and expectation for future work are provided in the conclusion.

DATA AND METHOD

The georeferenced data which concerns the incidents of property felonies in the city were extracted from Google Maps web service using a Java algorithm for refinement. The web feed is made by a Brazilian organization that collects reports from users whenever one's submitted to an offense or witness any property felony or misdemeanor.

Users may address to the website http://www.ondefuiroubado.com.br/ and rovide relevant information with regard the address where the crime occurred, the type of felony, date and time of incident, stolen objects, estimated loss in R$ (Brazilian Real) currency, and victim's genre. A total of 1199 incidents were obtained in this first data extraction, which concerns the period from August 2013 to June 2016, distributed along the urban space with particular concentrations in the southeast and center east regions, which also happens to be the wealthier and bustling spots in the city, as presented by the dark-red feature points on Figure 1.

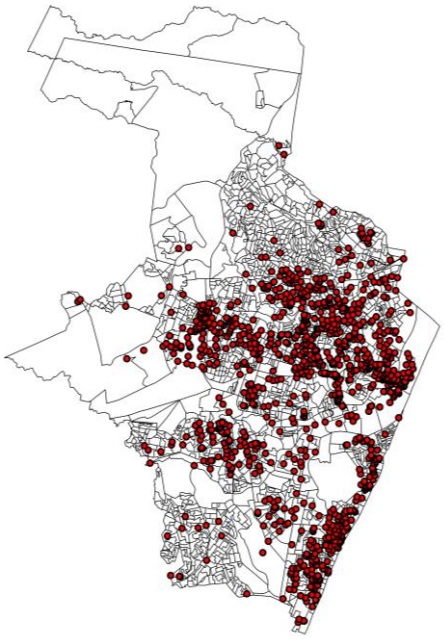


Figure 1 – Property Felonies' Distribution from 2013 to 2016

Recife has a population of 1.537.704 residents and about 218,435 km² urban and rural area decomposed into 1852 census sectors that put together regions with similar socio-economic characteristics. From the geospatial dataset, 11 types of property felonies are considered: Theft, Robbery, Robbery committed to a group of people, Express (Lighting) Kidnapping, Burglary (Housebreaking), Auto Burglary, Commercial Burglary, Saucy Bank, Motor Vehicle Robbery or Theft, Flashrob (Steaming) or Robberies in Public Transportation, and Attempted Robbery. For the present work, only 2 types of felonies (Flashrobs and Public Transportation Robberies) and Attempted Robbery, and 2 grouping classes (Vehicle-related Felonies and Street Robbery) are considered. The class 'Vehicle-related Felonies' clusters crimes of Vehicle Motor Robbery or Theft and Auto Burglary. As for 'Street Robbery', it clusters an instance of felony committed in the streets by one offender, i.e. Thefts or Larceny, Mugging, Robbery and Robbery committed to a group of people.

The Optimized Hot (and Cold) Spots identification based on the [10] spatial autocorrelation statistics is adopted as the method to recognize criminal patterns along the urban space where the feature points that represent the property felonies are distributed. The validation of this methodology results a Gi* statistic for each feature point that represents a z-score for one property crime indent in the available layer, and an associated p-value which will define the grade of clustering association (whether it will represent a hot spot, cold spot, or radon pattern undefined clustering). The [11] Gi*

statistic is given by the Gi statistic minus its expectation divided by the variance square root:

$$G_i^* = \frac{\sum_{j=1}^{n}(w_{i,j}x_j) - \bar{x}\sum_{j=1}^{n}(w_{i,j})}{S\sqrt{\frac{n\sum_{j=1}^{n}(w_{i,j})^2 - (\sum_{j=1}^{n} w_{i,j})^2}{n-1}}} \quad (1)$$

Where $w_{i,j}$ and x_j are the spatial weight between the neighborhoods (census sectors) i and j and the attribute for feature i, respectively. The Gi* statistic provides the information of where neighborhoods with either high (depictured by red hot spots) or low (depictured by blue cold spots) incidents cluster spatially. To be statistically significant, i.e. to present a considerable spatial association, the sum of incidents from each census sectors i and its neighboring sector j is compared to what should be an expected value of incidents, taking into consideration the sum of all sectors (see equation 1). When the difference from the sum of incidents of i and j with regard the expected value of the sum of incidents is large enough to reject the hypothesis the both are statistically similar, then a hot spot (a sector with high occurrences surrounded by other sector with high incidence) or cold spot (sector with low incidents surrounded by other sector with low incidence) will be designed.

The importance of this methodology in the visualization of georeferenced data leads to additional arguments to support better decision making of public resources allocation, once the inference of policing strategies is not exclusive based on the quantity of occurrences. Some assessed regions with a large amount of property crimes incidents might not be enough to be considered as a statistically significant hot spot. For this purpose, beyond the high value of the set of incident, it remains necessary that the incidents be surrounded by other high value incidents, which in Figure 1 is represented by the dark-red feature points. Then, the spatial association is corrected for false-positive differences using [8] False Discovery Rate Control, which results statistically more significant z-score for the spatial evaluation.

SPATIAL ASSESSMENT

The spatial pattern observed by the red-to-blue scale in Figures 2 to 5 measures the strength of clustering concentration as the high and low incidences of the related felony are placed on a given census region. The feature points of property crime incidents presented in Figure 1 were aggregated into 1852 polygons that represent the census sectors of Recife for this analysis. The hot and cold clusters brings statistical evidences of an underlying spatial pattern based on the assumed confidence level the z-score information provided to each aggregated feature. Positive z-scores for a p-value below its assumed confidence level brings empirical evidences for a significant hot spot, where high incidents of the related felony cluster together, and negative z-scores for a p-value below its assumed confidence level indicates a significant cold spot where low incidents of the analyzed felony cluster together. For the optimized clustering, the Inverse Distance approach is applied which implies that closer features are weighed more heavily than features that are distant

The scale of analysis is defined by incremental spatial autocorrelation performed with support of Moran' I statistic [9], measuring the intensity of spatial clustering in seven dimensions, 0.01, 0.05 and 0.1 confidence levels for both high and low

incidences' clustering, and a non-significant scale defined whenever the aggregated feature layer does not present enough statistical evidence for an accurate relation of the studied felony with the environment. For instance, the hot spot that clusters the thefts and robberies in public transportation observed in the second quadrant of Figure 2 is composed by 12 census sectors (polygons) which belong to the neighborhoods of *Várzea* and *Cidade Universitária*. The highest positive z-score is valued by 7.3731 for a p-value equals to 0 and the lowest positive (in orange scale) has z-score valued by 3.1479 for a p-value 0.001644. A set of spatial relationship can be constructed from the data with the environment technical features or socio-economic and demographic characteristics, which might support the better allocation of public resources, technologies and policing strategies.

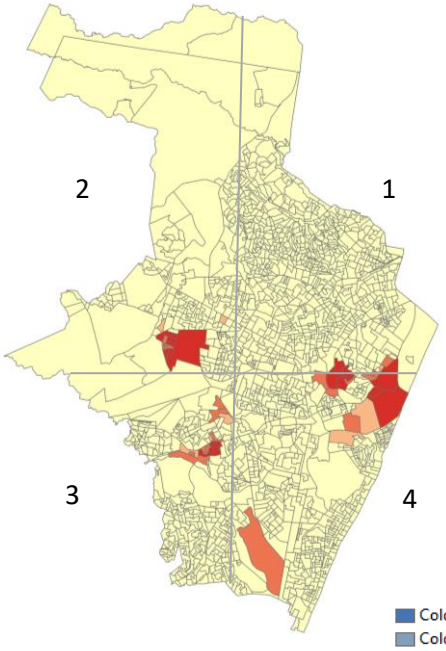


Figure 2 – Public Transportation

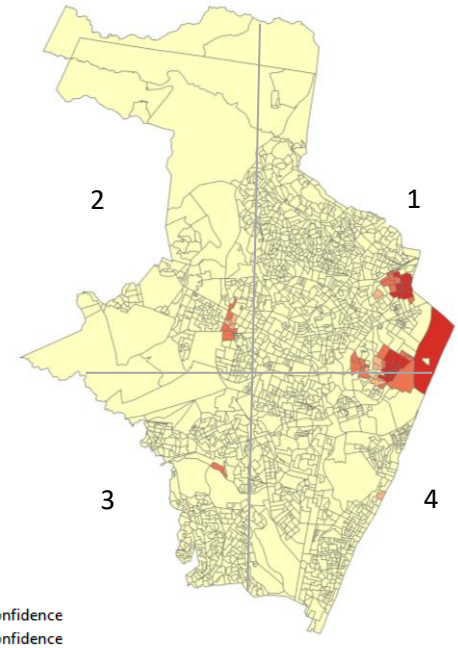
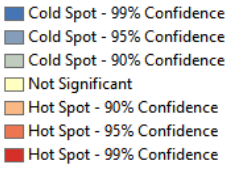
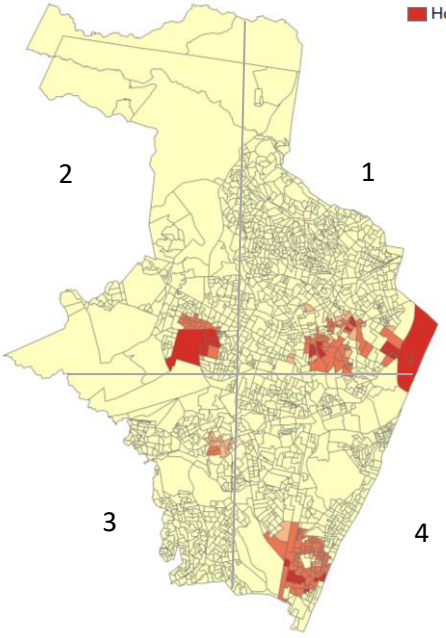


Figure 3 – Attempted Robbery





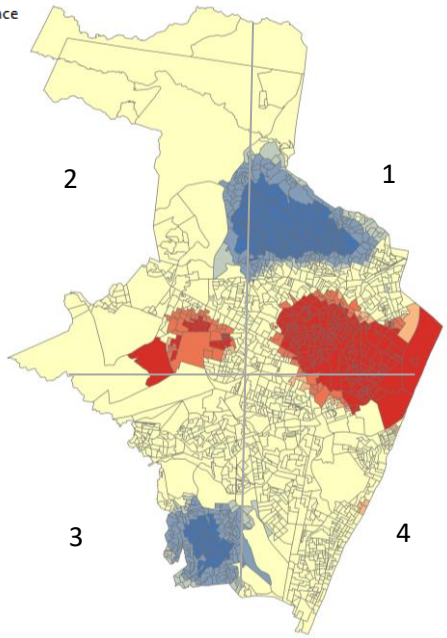


Figure 4 – Vehicle-related Felonies

Figure 5 – Street Robbery

CONCLUSIONS

The spatial visualization of property felonies in this work followed both a systematic and a subjective data curation process that has implication over the determination of public safety strategies to investigate the criminal behavior distributed on the urban space. The systematic process is made with the collection, and heuristic to organize and integrate the spatial references provided by the web-site big data repository of property offenses. The support of *Gi** statistical tool makes feasible the construction of optimized clusters, in which significant information regarding the concentration and dispersive tendencies over the analyzed census sector can be extracted and evaluated.

The subjective side of the curation process requires substantial experience from the analyst to understand the grade of discrimination in which each felony might reach in the disaggregation procedure. In the present work, this has been done by a meticulous checking on the motivation and expectations underling the each criminal conduct in our data. Certainly, the subjective nature of our assessment, in special, undergoes for several barriers when the analyst judgment over the criminal patter causes a misleading perception of the observed reality by an inappropriate discrimination of some types of felonies, or an inaccurate aggregation of others.

Another limitation resides on the quality of the collected data. Once it regards victimization data, provided by any individual type connected to the web network, they might be subjected to errors or wrong bias. Nevertheless, this concern does not undermine the importance of geoinformation systems to support the decision making of public policy makers. In addition, a tight data mining process might be able to minimize this issue. Finally, the next steps will rely on exploratory regressions and further investigations on the interaction that each disaggregated criminal conduct described by the cluster's feature layers has on the potential spatial, demographic, social and economic determinants of criminality. Much of discussion is still under construction, and good prospects for the public safety decision makers is expected.

REFERENCES

1. Bernasco, W., & Elffers, H. (2010). Statistical analysis of spatial crime data. In Handbook of quantitative criminology (pp. 699-724). Springer New York.

2. Chainey, S., Tompson, L., & Uhlig, S. (2008). The utility of hotspot mapping for predicting spatial patterns of crime. Security Journal, 21(1-2), 4-28.

3. Tabangin, D. R., Flores, J. C., & Emperador, N. F. (2010). Investigating crime hotspot places and their implication to urban environmental design: a geographic visualization and data mining approach. International Journal of Human and Social Sciences, 5(4), 210-218.

4. Sherman, L. W., Gartin, P., & Buerger, M. E. (1989). Hot spots of predatory crime: Routine activities and the criminology of place. Criminology, 27, 27-55.

5. Nasar, J. L., & Fisher, B. (1993). 'Hot spots' of fear and crime: A multi-method investigation. Journal of environmental psychology, 13(3), 187-206.

6. Chainey, S., & Ratcliffe, J. (2013). GIS and crime mapping. John Wiley & Sons.

7. Maltz, M., Gordon, A. C., & Friedman, W. (1991). Mapping crime in its community setting: Event geography analysis. Michael Maltz.

8. Caldas de Castro, Marcia, and Burton H. Singer (2006). "Controlling the False Discovery Rate: A New Application to Account for Multiple and Dependent Test in Local Statistics of Spatial Association." Geographical Analysis 38, pp 180-208.

9. Moran, P. A. P. (1950). "Notes on continuous stochastic phenomena." Biometrika, 37, 17–23.

10. Getis, A. and J.K. Ord. (1992). "The Analysis of Spatial Association by Use of Distance Statistics" in Geographical Analysis 24(3).

11. Ord, J.K. and A. Getis. (1995). "Local Spatial Autocorrelation Statistics: Distributional Issues and an Application" in Geographical Analysis 27(4).

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

ICDSST 2017 on Data, Information and Knowledge Visualisation in Decision Support Systems

Data Visualization and Visual Analytics for Mountain Rescue Service Ski Injury data: Mt. Kopaonik, Serbia case

Sandro Radovanović, Boris Delibašić, Milija Suknović
University of Belgrade, Faculty of Organization Sciences
Jove Ilica 154, 11000 Belgrade, Serbia
{sandro.radovanovic, boris.delibasic, milija.suknovic}@fon.bg.ac.rs
web-page: https://www.fon.bg.ac.rs

ABSTRACT

Development and use of information technologies resulted in the generation and accumulation of a large amount of data which are underused. In order to gain insight into such data collections, we can use of data visualization and visual analytics technologies. This paper discusses the usage of data visualization and visual analytics technologies on Serbian mountain rescue service data, i.e. ski injuries from Mt. Kopaonik, Serbia ski resort from 2005 to 2016. We propose a web-based application for interactive exploratory data analysis and temporal visualization of historical data to obtain useable knowledge, which can be used as a decision support system. Interactive data visualization and visual analytics allowed us to detect patterns that lead us to interesting insights in relation to space, time and specific types of injuries. We demonstrated the power of data visualization and visual analytics for the discovery of interesting patterns in data related increase of a number of injuries per year, injuries per time of day, and injury types and places. Limitation of this visual analysis is that data refers only to injured skiers, not to total skiing population. We note that the findings are supported in the literature. However, some findings are unique for Mt. Kopaonik, Serbia.

Keywords: Data Visualization, Visual Analytics, Ski Injuries

INTRODUCTION

It is noticed that data visualization and visual analytics enable decision makers' ability to understand the problem, the reason for it and make a decision on the daily basis. This is especially important in the era of Big Data [12, 6] and complex datasets [8]. Tools and techniques for data visualization and visual analytics are used to visually represent the data with the idea to detect some expected and unexpected relationships, deviations from predicted or expected values, and most importantly present information in an effective and understandable form. Data visualization and visual analytics are focused on analytical

reasoning with an idea to help decision-making process throughout the entire analytic process, not only for presentation or viewing [7, 19].

Winter sports and leisure are an industry which is interesting for researchers and other stakeholders namely because of high financial resources involved (in United States ski resorts in 2014/2015 there were 6.5 billion euros of direct spending). Additionally, there are 57.1 million skier visits (skier-days) reported since season 2002/2003 [15]. Having this in mind, and the fact that injury rate is 2 injuries per thousand skier days [13] ski injuries are frequent and present interesting research and ski resort problem. Data visualization and visual analytics for each ski resort are important as injury patterns can differ from resort to resort and from season to the season [5]. In this paper, we used Serbian mountain rescue service data which recorded ski injuries from Mt Kopaonik from 2005 to 2016. There are over 6,000 recorded injuries. However, this number is misleading because there are a lot more unreported injuries. Mt. Kopanik has a high number of skiers with over 100,000 skier days and over 1,000,000 skier slopes yearly [3].

Although ski injuries are not often costs are very high. One can twist a knee, dislocate a shoulder or even die. These injuries not only affect the individual but affect ski resort (less skiing, less utilization of resources, bad reputation) and overall society (one can be unable to work after injury temporary or permanently). In order to identify risk slopes and prevent injuries, one can utilize visual analytics which would provide better overall quality of service.

There are already numerous studies in skier's individual injury risk factor identification. Various risk factors have been reported like: gender, age [13], personality types [1], skier collision [4], skiing errors [2], speed of skiing [4, 2], fatigue [2], perception of low difficulty [2], skillfulness and experience [4], quality of equipment [4], quality of ski slopes and quality of their preparation, collision against objects, and jumps [4]. Beginners usually have an injury rate five times that of experts [10].

For the purpose of this paper, we developed a website with Tableau software [18], which allows visual analytics of ski injury data across more Mt. Kopaonik, Serbia ski resorts, and through eleven consecutive seasons.

MATERIALS AND METHODS

For the purpose of this paper, a website for data visualization is developed. It offers better insights into distribution and specific characteristics of ski injuries in Mt. Kopaonik, Serbia. The analysis was done on real Serbian Mountain rescue service actions data. The purpose is to demonstrate the idea of interactive exploratory data analysis and the capabilities of the data visualization and visual analytics for decision support systems and decision making. The interactive part of the website uses Tableau Online [18]**,** which are suited for interactive visualizations for business intelligence.

The underlying database that was used in experiments as well as in the prototype of visual data analytics was obtained from Serbian mountain rescue service team for a period of eleven years (2005 - 2016). It consists of 7425 rescue actions, from which 6569 are valid. From those actions where gender was entered, we observe that 3724 (56.69%) were males, while 2841 (43.25%) were females (gender of remaining are not identified). It is worth to observe that average age at the time of the injury (mean ± standard deviation) is 28.98 ± 15.08 years. More specifically, for male average age at the time of the injury (mean ± standard deviation) is 29.07 ± 14.97, and for female average age at the time of the injury (mean ± standard deviation) is 28.87 ± 15.19.

We will perform data visualization to present *global* distribution of the data values, and after that *local* distribution of the data values around a data point of interest [14]. This way we first observe high-level data characteristics such as trends, outliers, frequencies, distributions, and correlations which allow us to get "big picture" of the data. Then, we utilize low-level data characteristics in order to inspect high-level data characteristics. This can be understood as a drill down in visualizations.

RESULTS

The project website (available at http://odlucivanje.fon.bg.ac.rs/project/ski-injuries-serbia/) demonstrates the effectiveness of the data visualization and visual analytics that can be used by a wider range of population ranging from ski expert to more general public. This website allows users to visualize and reason some of the most interesting insights that can be offered using the available data. This includes an interactive dashboard of a number of ski injuries per gender and per location (Figure 1). We must emphasize that every dashboard present number of injuries and it is not normalized with a total number of skiers.

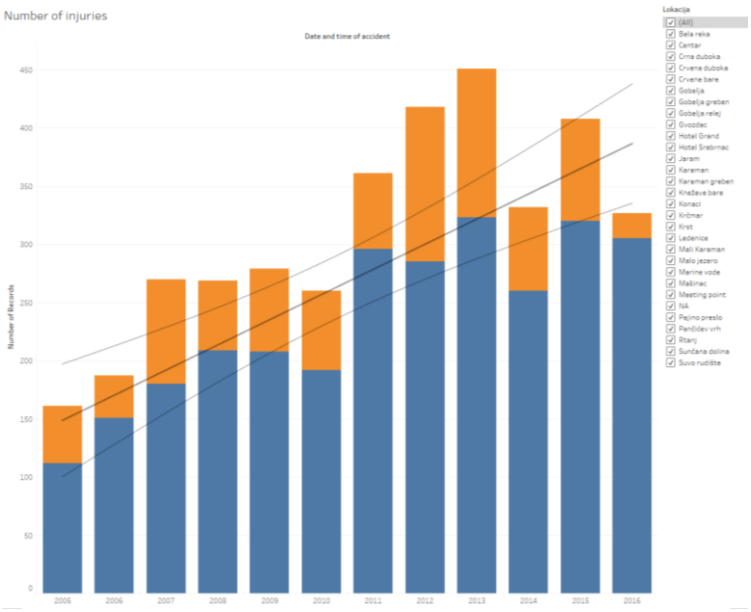


Figure 1: Dashboard displaying number of injuries per gender and location

First, we must state that this is not stacked bar chart, but bar chart with overlapping bars where blue color represents female skiers and orange represent male skiers. First, one can notice that overall trend is increasing. However, one should cautiously interpret this data, as the data has not be normalized with the number of ski slopes. Also, one can notice that male skiers experience more injuries compared to women in every year. This is probably, but cannot be concluded for sure, due to the fact that is more male skiers compared to female skiers. We emphasize that this dashboard has additional filter by the location of ski slopes where one can observe trends of injuries for every ski slope. We can observe that there are several ski slopes that have small number of injuries (< 10 in total), or stable number injuries (number does not change) or even decreasing number of injuries over time, but also we observe several ski slopes which does have increasing number of injuries such as Krst (1.14 times more injuries per year), Crna duboka (1.86 times more injuries per year), Mali Karaman

(1.97 times more injuries per year), Crvena duboka (2.75 times more injuries per year), Pančićev vrh (3.03 times more injuries per year), Karaman greben (4.04 times more injuries per year) etc.

Further, we wanted to inspect age of skier and injury (Figure 2). One can notice that number of injuries are most frequent from 11:00 to 13:00 hours and that average age is around 29 years. Also, one can notice that help time is fairly constant at five minutes. However, using this dashboard we can find interesting patterns such as fact that age at injury for Bela reka and Krčmar are higher than average and that help time is reducing over years (around 17% and 20% per year, respectively), or that ski slope Gobelja and Gobelja relej have younger skiers if we look at age at injury. However, on a ski slope, Karaman one can observe an increase in help time (around 7% per year).

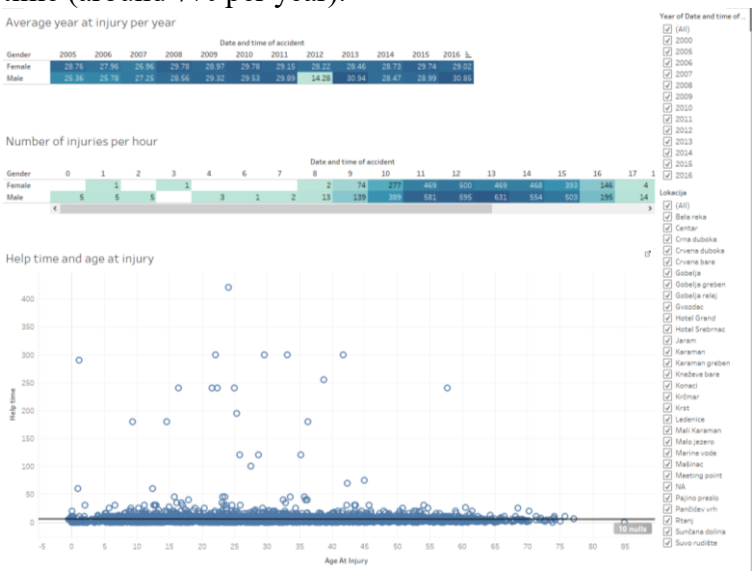


Figure 2: Dashboard displaying average age, number per hour of day and help time per gender and location

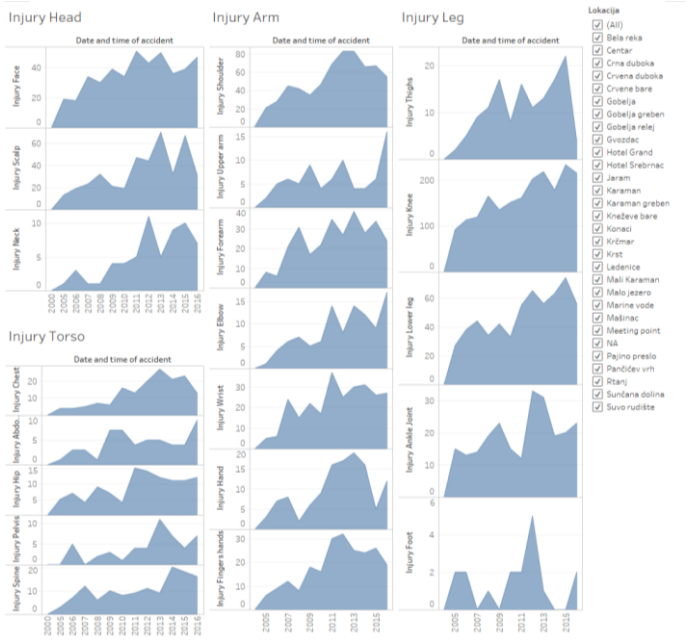


Figure 3: Dashboard for visual inspection and analysis of injuries

Finally, we created a dashboard for visual inspection and analysis of injuries (Figure 3). We divided dashboard by a grouping of injuries. One can observe that there are most injuries of knees, then shoulders, scalp, lower leg, and face. Further inspection showed that Karaman greben and Mali Karaman slopes are especially hard for knees since most of the knee and lower leg injuries are on those slopes. However, some ski slopes do not have many injuries of any types. This probably means that those ski slopes are underused. This information can be used by the decision maker in order to balance skiers and make ski resort less crowded.

DISCUSSION AND CONCLUSIONS

Findings obtained with this visual analytics and data visualization website are in accordance with finding on other ski resorts. Namely, knee injuries, which are a most common injury in Mt. Kopaonik, Serbia, are also most common in other ski resorts [16, 11]. Also, head injuries, although not most frequent are a major risk in ski industry accounting for approximately 18–27% of all injuries [17]. Similar findings are observed in Mt. Kopaonik, Serbia. The reason why ski injuries are that common can be found in the literature [4, 2, 13, 17**]** which emphasizes factors like sensation seeking, high impulsivity, and skiing proficiency is associated with high-risk behaviors. However, all finding, similarities, and dissimilarities between ski resorts have to be further inspected. Additionally, we must state that Mt. Kopanik, Serbia is fairly similar to other ski resorts with 0.2 ski injuries per 1000 skier-days [3] which is stated also in [13].

Based on these dashboards decision maker can create health care strategy and see demography of injuries. In the example, rescue service can provide more members at specific time frame and specific location in order to prevent injuries from happening. Some slopes are known as risk and one can monitor statistics. Further, if we combine this data with ski resort data we can create real-time dashboards and calculate the risk of injury for each skier which would provide decision support system for every stakeholder.

Although this paper uses only ski injuries data from one ski resort, we developed a solution that can be applied to every ski resort and which can cover different types of injuries and present them in real time. We are aware that our data is limited and can be enriched with i.e. a total number of skiers or hospitalization data in order to obtain better and more complete insights further leading to better decision making. Furthermore, we expect the results of this project to be used in our further work on ski injury prediction and optimization of ski service resources. Interpretation of the data and results demonstrated and obtained from this paper can be used as a supplement to further research.

REFERENCES

1. Castanier, C., Scanff, C. L., & Woodman, T. (2010). Who takes risks in high-risk sports? A typological personality approach. *Research quarterly for exercise and sport*, *81*(4), 478-484.
2. Chamarro, A., & Fernández-Castro, J. (2009). The perception of causes of accidents in mountain sports: a study based on the experiences of victims. *Accident Analysis & Prevention*, *41*(1), 197-201.

3. Delibašić, B., Marković, P., Delias, P., & Obradović, Z. (2016). Mining Skier Transportation Patterns From Ski Resort Lift Usage Data. *IEEE Transactions on Human-Machine Systems*.
4. Dohin, B., & Kohler, R. (2008). Traumatologie du ski et du snowboard chez l'enfant et l'adolescent: épidémiologie, physiopathologie, prévention et principales lésions. *Archives de pédiatrie*, *15*(11), 1717-1723.
5. Greve, M. W., Young, D. J., Goss, A. L., & Degutis, L. C. (2009). Skiing and snowboarding head injuries in 2 areas of the United States. *Wilderness and Environmental Medicine*, *20*(3), 234-238.
6. Janvrin, D. J., Raschke, R. L., & Dilla, W. N. (2014). Making sense of complex data using interactive data visualization. *Journal of Accounting Education*, *32*(4), 31-48.
7. Keim, D. A. (2002). Information visualization and visual data mining. *IEEE transactions on Visualization and Computer Graphics*, *8*(1), 1-8.
8. Keim, D., Andrienko, G., Fekete, J. D., Görg, C., Kohlhammer, J., & Melançon, G. (2008). Visual analytics: Definition, process, and challenges. In *Information visualization* (pp. 154-175). Springer Berlin Heidelberg.
9. Konik, I. G., Kermanc, M. H., Mentis, A. F. A., Germeni, E., & Petridou, E. (2010). Fall-related head injuries among alpine skiers and snowboarders in Slovenia. *Injury Prevention*, *16*(Suppl 1), A208-A208.
10. Laporte, J. D., Bajolle, L., Lamy, D., & Delay, J. B. (2012). Winter sports injuries in France over two decades. In *Skiing Trauma and Safety: 19th Volume*. ASTM International.
11. Mayr, H. O., & Chèze, L. (2016, May). The Knee Injuries in Skiing: Mechanisms and Assessment. In *ISBS-Conference Proceedings Archive* (Vol. 33, No. 1).
12. Miller, H. G., & Mork, P. (2013). From data to decisions: a value chain for big data. *IT Professional*, *15*(1), 57-59.
13. Ruedl, G., Kopp, M., Sommersacher, R., Woldrich, T., & Burtscher, M. (2013). Factors associated with injuries occurred on slope intersections and in snow parks compared to on-slope injuries. *Accident Analysis & Prevention*, *50*, 1221-1225.
14. Schulz, H. J., Nocke, T., Heitzler, M., & Schumann, H. (2013). A design space of visualization tasks. *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2366-2375.
15. Shealy, J., Ettlinger, C., Scher, I., & Johnson, R. (2015). 2010/2011 NSAA 10-year interval injury study. In *Skiing Trauma and Safety: 20th Volume*. ASTM International.
16. Shulman, R., Teng, L., & Salonen, D. (2016). Imaging of Snow Skiing and Snowboarding Injuries. In *Imaging in Sports-Specific Musculoskeletal Injuries* (pp. 381-402). Springer International Publishing.
17. Thomson, C. J., & Carlson, S. R. (2015). Increased patterns of risky behaviours among helmet wearers in skiing and snowboarding. *Accident Analysis & Prevention*, *75*, 179-183.
18. Vidhya, S. S. N. S. S., Sarumathi, S., & Shanthi, N. (2014). Comparative Analysis of Diverse Collection of Big Data Analytics Tools. *World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering*, *8*(9), 1646-1652.
19. Zhang, L., Stoffel, A., Behrisch, M., Mittelstadt, S., Schreck, T., Pompl, R., ... & Keim, D. (2012, October). Visual analytics for the big data era—A comparative review of state-of-the-art commercial systems. In *Visual Analytics Science and Technology (VAST), 2012 IEEE Conference on* (pp. 173-182). IEEE.

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

A DSS for a Bank Risk Analysis

Rakotoarivelo Jean Baptiste[1,2], Zaraté Pascale[3]
Viguié Philippe[4], Kilgour D. Marc[5], Michel Février[4]

1 : IRIT – Toulouse University - 118 route de Narbonne 31062 Toulouse Cedex 9 – France
2 : Université de Toamasina - Sis 5, rue Bir Hakeim, Camp Manguier BP : 591 – Toamasina 501 – Madagascar, Jean-Baptiste.Rakotoarivelo@irit.fr ; rjbravelo@yahoo.fr
3 : IRIT – Toulouse Capitole University - 2 rue du Doyen Gabriel Marty 31042 Toulouse Cedex 9 – France , Pascale.Zarate@irit.fr
4: CEMP, Toulouse, France
philippe.viguie@cemp.caisse-epargne.fr, Michel.Fevrier@cemp.caisse-epargne.fr
5: University Wilfrid Laurier, Waterloo - 75 University Avenue West, Waterloo, Ontario N2L 3C5 Canada, mkilgour@wlu.ca

Abstract. This article was designed to analyze different policy securing information inherent in bank or financial institutions information systems. The work is based on the development of DSS for the identification and prioritization of risks for Information System. The goal of this system is to support a decision maker by reducing the risks and intrinsic errors in the information related to the financial sector activities system. The proposed DSS is based and the AHP[1] methodology. This method supports each manager to the choice of a specific decision in the development of management to demonstrate the organizational capacity to significantly reduce the losses associated with security vulnerabilities of information systems.

Keywords: banking risks, information systems, bank security, DSS, AHP.

1 INTRODUCTION

Security and privacy are two very important aspects in the areas of financial institutions and banks; they determine its own security policy to ensure the safety means for their businesses as personal and property. According to [1], we firstly specify the security objectives of an information system of a bank in a meaningful way, and secondly, we develop a DSS based on the AHP methodology to support a decision maker in order ensure a certain level of confidence in the security of the system.

1.1 Problem statement

This work is based on the analysis of the Information System in a real bank. This analysis is possible by the definition of 4 criteria:
1. **Disponibility**: The ability of the information system to guarantee processing performance and access to information under predefined conditions, the unavailability of the information system occurs in the case of: a transient or repetitive disruption of the business asset or

[1] Analytic Hierarchy Process : AHP

associated services or a long unavailability of the associated business or service asset at a critical deadline for the business.
2. **Integrity:** The ability of the information system to ensure that the information is unalterable in time and space, loss of integrity occurs in the case of: loss or alteration of data or use deviated from processing under a Fraud or malicious use or modification of treatment due to an anomaly or human error in the data entry or use of treatments.
3. **Confidentiality**: The ability of the information system to protect sensitive information from unauthorized disclosure, a loss of confidentiality occurs in the event of: breach of confidentiality in the chain of processing or procedures or malicious or fraudulent access to Data.
4. **Proof and Control**: The ability of the information system to provide audit trails and the evidence corresponding to the actions performed, a lack of evidence appears in the case of: a denial of action, loss or absence of evidence.

These are the two types of business impacts which are defined as: the financial impact can be direct or indirect result of various financial failures (loss of market share, penalties, damages and interest) and the image is the impact degradation of the image by reporting to different types of customers in connection with this system. These four criteria are used to reduce the risk on two axises: financial risk and image risk. For example in Table 1, D(8,3) means that on the criteria Disponibility the solution 1 has 8 for the financial risk and 3 for the image risk. Five possible solutions are then possible for decision makers that are called "Solution". Each alternative is then marked on each criterion for the two view points: Financial and Image. These scenarios of these treatments are represented in Table 1 below

Table-1: Alternatives Evaluation

Solutions	Disponibility	Integrity	Confidentiality	Proof & Control
Solution 1	D(2,2)	I(3,2)	C(4,1)	P (1,2)
Solution 2	D(8,3)	I(3,2)	C(6,0)	P (7,5)
Solution 3	D(1,2)	I(3,1)	C(2,1)	P (4,3)
Solution 4	D(11,3)	I(2,4)	C(11,2)	P (3,2)
Solution 5	D(2,1)	I(3,3)	C(3,1)	P (1,0)

The problem statement is represented by two figures 1 and 2. The model of the problem to solve is the following:
Level 0: the main goal of treatment: classification profile of the business activity on one point of view;
Level 1: the four criteria Disponibility (D) Integrity (I), Confidentiality (C), Proof and Control (P);
Level 2: the five "Solutions"
Level 3: Alternatives: Financial Impact, Image Impact

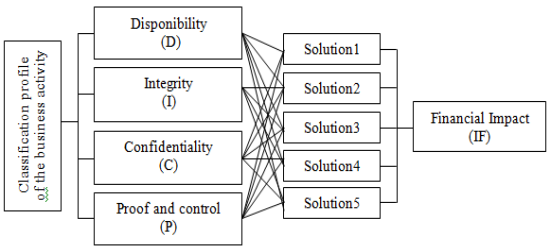


Figure 1: Hierarchical decomposition profile: Financial Impact.

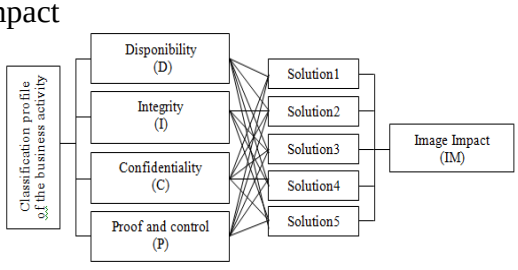


Figure 2: Hierarchical decomposition profile: Image Impact

2 PROPOSED METHODOLOGY

The proposed Decision Support System is an application of the AHP method which is an acronym of Analytic Hierarchy Process. This methodology is a multi-criteria methodology characterized by the determination of the weights of the criteria and alternatives. The approach is possible by a pairwise comparison at each level of the hierarchical problem. In a first step we want to minimize all risks and we decided to have all criteria having the same weight. It is the reason why each criteria compared to another has a "-1" mark. Table 2 shows the comparison of criteria.

Table-2: Comparison of criteria: Parameter value

Comparison of criteria	Disponibility	Integrity	Confidentiality	Proof & Control	Priority
Disponibility	-1	-1	-1	-1	0,25
Integrity	-1	-1	-1	-1	0,25
Confidentiality	-1	-1	-1	-1	0,25
Proof & Control	-1	-1	-1	-1	0,25

2.1 Different treatment steps

The treatment is started with a sheet " parameter " of the comparison criteria shown in Table 2 above, and at the end of treatment we weighted the results of each criteria (D), (I), (C), (P) by priority vector value in table 2 [3], [4], [5], [6], the weighting trick is called aggregation [7] We took the initial value shown in table1 for the Financial impact (IF), the Image impact (IM).

Step 1: Profile classification of business of all processes (IF) and (IM)
In this case we compared two by two values for each active business in table 3 and 4 below. The results of comparisons provide a first result that will be used in the following process.

Table-3: Profile classification of the business assets of all processes: (IF)

Financial Impact	Disponibility	Integrity	Confidentiality	Proof & Control
Solution 1	-2	-3	-4	-1
Solution 2	-8	-3	-6	-7
Solution 3	-1	-3	-2	-4
Solution 4	-11	-2	-11	-3
Solution 5	-2	-3	-3	-1
Total	-24	-14	-26	-16

Table-4: Profile classification of the business of all processes: (IM)

Image Impact	Disponibility	Integrity	Confidentiality	Proof & Control
Solution 1	-2	-2	-1	-2
Solution 2	-3	-2	0	-5
Solution 3	-2	-1	-1	-3
Solution 4	-3	-4	-2	-2
Solution 5	-1	-3	-1	0
Total	-11	-12	-5	-12

Step 2: Calculation method

We will proceed to the calculation for each criterion (D, I, C, P) of five "solutions". In this case the diagonal of this matrix is equal to "1 ". The method of calculation is done by comparison in adding (because it is the minimization)

Steps-3: Determination of weight criteria
According to the formula above, we have the results of the weight of four financial impact criteria same procedure for calculating the weight of four images impact criteria.

Steps-4: Parameter value and aggregation
After the calculations, we have the combination mode is called "total aggregation" and the value 0, 25 included in each table is the weight parameter shown in the previous table 2. This value is used to determine the total aggregation.

Steps-5: Total aggregation

These results are summarized in Table 5 and Table 6. The way of grouping is called Approach total aggregation (weight criteria are multiplied by the parameter value)

Table 5: Total weighted aggregation 100: Financial Impact

	D	I	C	P
Parameter value	**0,25**	**0,25**	**0,25**	**0,25**
Solution1	4,38	5,06	4,72	4,38
Solution2	5,57	5,06	5,07	6,26
Solution3	4,27	5,06	4,46	5,14
Solution4	6,38	4,76	6,18	4,83
Solution5	4,38	5,06	4,58	4,83

Table 6: Total weighted aggregation 100: Image Impact

	D	I	C	P
Parameter value	**0,25**	**0,25**	**0,25**	**0,25**
Solution1	4,89	4,81	4,90	4,74
Solution2	5,33	4,81	4,28	6,13
Solution3	4,89	4,52	4,90	5,14
Solution4	5,33	5,65	6,02	4,74
Solution5	4,56	5,20	4,90	4,26

2.2 DSS Implementation

Modeling preference of the decision maker is already presented in Table 1 above, we have exploited the information provided by the decision maker and supported by the AHP method, and after different modes of treatment [8], [9]. We have two different results: (1): The initial value is the result of weighting priority of the parameter value (0,25) by the aggregation of tables 5 and 6. In this phase all the criteria values are equal to the financial impact and the impact image. Second result (2): the final value is the result of the weight criteria (D, I, C, P) of the financial impact aggregation and image impact weighted by the value given by the responsible factor. The coefficient is the scale value of each criterion from "**0**" to "**n**". In this phase we adopt the basic mechanism is the pairwise comparison of the shares each criterion [10] .The actions: are the business assets already defined by the organization's decision-maker Criteria: the decision parameters based on defined value shares in a totally ordered, representing an objective decision maker, these are the criteria (D, I, C, P).

3 DSS REVISION TO A NEW APPLICATION

The stage of our research is advanced by the finding of Chief Information Security Officer (CISO), according to the maximization by the AHP method via the means for minimizing reverse the finding rests on the means of verification in order to have an idea says the decision to be taken at the previous processing results. We need to spread a new application to easily meet the need of SSI responsible, these applications are composed of 86 ways of arrangement in which the input of a value for each criterion may respond directly to an exact answer formalized by the CISO.

3.1 Different treatment steps

Steps-1: Procedures calculations

In this new application we retain the criterion of comparison scale equal to "-1" Table 2 "Parameter value" and a "-" for comparison values per pair of identifying application. We compared two by two alternative 86 named "Application ID".

Steps-2: calculation method

Comparison criteria (D, I, C , P) , the values of priorities for each comparison are weighted by the weight of four criteria in the initial result and the final result of the aggregation total , early values are all equal to 1 and priorities are the same values equal to 0,25. We compare two by two each criterion such as in the case of Availability.

Steps-3: Initial value

This value is the arrangement of scale of data proposed by the CISO, it matches the number of identities of 86 applications, throughout treatment, are all negative initial value of each criterion.

Steps-4: square matrix calculation 86 x 86 criteria (D, I, C, P)
For method AHP the diagonal of this matrix is equal to "1 " in this treatment method is calculated by comparing by adding (as it is the minimization) away from the first line and the second below, and the first line with the third row and so on down to 86th lines, the value of goodwill is divided by the sum of criteria relating to the identity of 86 applications., the results obtained are registered online, and reverse are placed in vertical columns.

Steps-5: Total aggregation
The results are used in the following calculation:
(1) Weighting the results of each criterion (D), (I), (C), (P) by the table2 parameter value, the result is called “Initial value”
(2) The "final value" according to the criterion of each entry of scale (D), (I), (C), (P) obtained by the weighting vector priority. The complete aggregation is to reduce in one way or another all the criteria to be considered by a single criterion.

3.2 DSS new implementation

We have two graphic presentations illustrating the processing result of this work:
(1): the set of graphical presentation of the initial value and the final value (Figure 3). In this graph the initial value remains unchangeable while the final value is varied according to the scale value to grasp and a corresponding indication to this value appears in the language of each criterion.

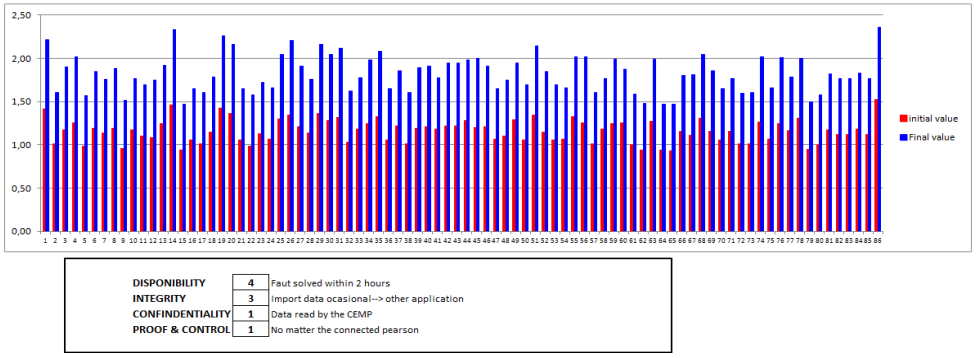


Figure 3: Scale value input screen and the graphic result of the initial value and the final value.

The graphic variation is dependent on the scale entry of the relative value for each case of criterion, and then each entry corresponds to an indication from the head of the SSI in order to easily remember their labels to when the graph examinations. Each of the abscissa value point consists of two histograms initial value and final value) and the ordinate axis is the value of weight criteria.

CONCLUSION

The treatment is evolved as the case concerning the need for the Chief Information Security Officer, this development is a three steps, the first is the processing means by maximizing method, the second is by reversing the first minimization method and the third is the application of a new proposal CISO by using an identity of 86 applications, and we keep looking until we have a better decision support system helping to improved information system for financial risk. In any case these treatments are framed in the following two distinct profiles: (1) the classification of the asset profile of all business process in which we determine two types of business impacts such as the financial impact and impact image. (2) The active profile support to operate the business assets having two types of classification

profile of impacts such as the financial impact and image. The results exhibit that within this organization, the ISS is increasingly addressed using approaches based on risk. The concept of risk, which remains intangible, is difficult to grasp: "The risks designate a future that is to prevent future". According to [10] risks related to the security must rest on techniques and specific methodologies. That is why we have applied the specific technique of the methods of multi-criteria decision support AHP to have a reliable result. Nevertheless, even modifying the criteria weights, this methodology does not allow to explicitly show a very important risk in comparison to others.

REFERENCES

[1] (SGDN/DCSSI/SDO/BCS), "Guide pour l'élaboration d'une politique de sécurité de système d'information," 51 Boulevard de la tour Maubourg 75700 Paris 07 SP, Mar. 2004.

[2] Y. Bouzarour Amokrane, "Structuration des Processus d'Aide à la Décision par Analyse Bipolaire," Thèse de doctorat, EDSYS Industriel 4200046, UT3 Paul Sabatier, Toulouse, 2013.

[3] M. Lauret Barut, N. Rouillé, M. Sanchez, "L'impact du nouveau paradigme réglementaire sur le rôle des banques dans le financement de l'économie," *Banq. Fr. Rev. Stabilité Financ.*, no. 19, p. 2, Avril 2015.

[4] A. Anwar, A. Gulzar, F. Bin Sohail, S. Naeem Akram, "in Impact of brand image, trust and affection consumer brand extension attitude: The mediating role of brand loyalty," *International journal of Economics and Management Sciences*, pp. 73–79, 2011.

[5] D.A Aaker, "Managing brand equity: Capitalising on the value of a brand name," *The free Press*, New York., 1991.

[6] A. Schârlig, "Décider sur plusieurs critères," *Panorama de l'aide à la décision multicritère, Série: Diriger l'entreprise*, Université de Lausanne, p. 304, 1985.

[7] J-M., Martel, Aggregating preferences: utility function and outranking approaches, University Laval, 1993

[8] V. Chankong, and Haimes, Y.Y., Multiobjective decision making theory and methodology, North-Holland, 1983, 406 pages.

[9] J. P. Brans, Ph. Vinke, *A preference ranking organization method: the PROMETHEE method*, Management Science. 1985.

[10] U. Beck, *La Société du risque- sur la voie d'une autre modernité*, Flammarion-Champs. 2003

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

Relative Importance Under Low-rank Condition and Its Applications to Semiconductor Yield Analysis

Zixin Shen, Argon Chen

Institute of Industrial Engineering, National Taiwan University
1, Section 4, Roosevelt Road, Taipei, 106, Taiwan
F01546034@ntu.edu.tw, achen@ntu.edu.tw
http://www.ie.ntu.edu.tw/dr_chen

ABSTRACT

In semiconductor manufacturing, multiple linear regression analysis is widely used to evaluate how a response is related to a set of potential factors. Since the yield loss is costly, it is of great interest to assess the "relative importance" of a factor by comparing the contributions made by individual factors in a particular regression model. This is difficult to determine, however, in the presence of correlated factors. To deal with this problem, dominance index analysis and relative weight analysis were proposed. Both indices offer a good measurement of relative importance even when factors are highly correlated. Relative weight analysis, however, involves significantly less computational burden than the dominance index when the number of factors in the final model is large. This makes the relative weight more useful as the high-dimensional data are more and more commonly observed. However, the relative weight analysis postulates that the design matrix should be full column rank, and is unsolvable when the sample size is smaller than the potential factors. This precondition limits the relative weight analysis from applications to industry cases where the production ramp-up resource is restricted and the manufacturing processes are highly complex with intertwined factor relations.

In this research, we extend the relative weight analysis to cases with the following two conditions: 1. a singular design matrix with a sample size greater than the number of factors, and 2. a sample size smaller than the number of factors. An actual semiconductor yield analysis case is used to illustrate the method and its effectiveness in detection of critical factors.

Keywords: Semiconductor, Multiple linear regression, Relative importance, Relative weight, Low rank, Singularity

INTRODUCTION

Semiconductor manufacturing consists of complex fabrication steps and has a long cycle time. When wafers have run through all the process steps, electrical test (ET) are performed to ensure whether each die on the wafer is functional. Since the semiconductor yield loss is not only costly but also affected by many potential factors, it is critical to discover association between an electrical testing (ET) parameter and multiple in-line process metrology items, and then find out the important factors causing the failed ET parameter.

Multiple linear regression analysis is commonly used to model the above association and identify the critical factors. Consider the following multiple linear regression model,

$$Y_{n*1} = X_{n*p}\beta_{p*1} + \varepsilon_{n*1} \quad (1)$$

where n is the sample size, p is the number of independent variables, Y_{n*1} is the dependent variable, X_{n*p} is an n-by-p design matrix, β_{p*1} is a p-by-1 vector of unknown regression coefficients, and ε_{n*1} is a p-by-1 vector of i.i.d random variables with zero mean and finite variance. To draw conclusions about the relative importance of the independent variables, researchers often examine the regression coefficients or the zero-order correlations with the dependent variable. When independent variables are uncorrelated, zero-order correlation between each independent variable and the dependent variable is equivalent to corresponding standardized regression coefficient. The squares of these p standardized regression coefficients sum to R^2, so the relative importance of each independent variable is the corresponding standardized regression coefficient, which can be expressed as the proportionate contribution each independent variable makes to R^2. However, these indices have long been considered inadequate in the presence of correlated independent variables [1]. Not only do the squared standardized regression coefficients no longer summed to R^2, but also two independent variables, that are very similar in nature with high correlation and should have comparable relative importance, may have very different regression coefficients.

When the independent variables are correlated, the topic of relative importance of independent variables is resurrected by the dominance index analysis introduced by Budescu in 1993[1]. The dominance index reflects the relative importance by the average usefulness of an independent variable when it is included with each possible combination of independent variables. The sum of dominance indexes over all independent variables is also equal to full model R^2. However, dominance index analysis requires the computation of R^2 for all possible $(2^p - 1)$ sub-models, which makes it very difficult and time-consuming to compute, especially as the number of independent variables increases. Realizing the limitations of dominance index analysis, a heuristic method called relative weight analysis is then proposed [2]. Relative weight analysis not only requires a much lower computation loading but also be shown to be highly consistent with dominance index through real case validation [2], and Monte Carlo experiments [3]. Relative weight analysis considers the relative importance for a set of independent variables under collinearity by using a set of uncorrelated variables that are most alike to the original variables. Relative weight analysis has two prerequisites, one is a sample size n not smaller than the number of independent variables p, the other is the observed independent variables not linearly dependent to one another. In other words, the design matrix X_{n*p} should be full column rank, and we refer to this condition as "regular condition". However, for semiconductor manufacturing or other industry applications where the yield loss or production experiment is extremely expensive and makes it uneconomical to obtain a sufficient sample size. That is, the sample size n may be smaller than the potential factors. In

another situation, the complex manufacturing processes may have factors that are interrelated according to physical laws resulting in a design matrix without full column rank, referred to as a singular X_{n*p}.

This paper aims to extend the relative weight analysis to allow two low-rank conditions, i.e., a condition where X_{n*p} is singular with $n \geq p$ and a condition where $n < p$. Finally, we will validate the proposed index for the "$n < p$" condition through a real semiconductor yield analysis case.

RELATIVE WEIGHT ANALYSIS

Relative Weight Analysis Under Regular Condition

Suppose that X is a *n*-by-*p* normalized matrix with full column rank, relative weight analysis is first to find a best-fitting orthogonal approximation to X_{n*p}, denoted by Z [4]. The formulation is as follows:

$$\begin{aligned} &\text{Min: } \operatorname{trace}(X-Z)^T(X-Z) \\ &\text{s.t.} \quad Z = XT \;;\; Z^TZ = I_p. \end{aligned} \tag{2}$$

where T is *p*-by-*p* transformation matrix and I_p is a *p*-by-*p* identity matrix. To solve the solution of equation (2), we find the SVD of X,

$$X = USV^T \tag{3}$$

where U is *n*-by-*p* matrix of left singular vectors, V is *p*-by-*p* matrix orthogonal matrix of right singular vectors, and S is *p*-by-*p* diagonal matrix with the singular values on the diagonal. The solution of Z is then given by the work [4] as following:

$$Z = UV^T \tag{4}$$

This new set of orthonormal variables Z can then be used to generate regression coefficients by regressing the dependent variable Y on Z:

$$\hat{\beta}_z = \{\hat{\beta}_{Z_l}\} = (Z^TZ)^{-1}Z^TY = VU^TY \tag{5}$$

where $l = 1, \dots, p$, Y is assumed to be a *n*-by-1 normalized vector, and $\hat{\beta}_Z$ is the *p*-by-1 vector of estimated regression coefficients linking the orthonormal variables Z to the dependent variable Y. Because variables in Z are orthonormal, the *l*th squared regression coefficients $\hat{\beta}_{Z_l}{}^2$ represents the relative importance of *l*th variable in Z unambiguously. However, because the variables in Z are only approximations of X, the values in $\hat{\beta}_Z{}^2$ are not the relative importance of each variable in X. Thus, Z should be related back to X to better estimate the relative importance of original variables. This is done by regressing X on Z:

$$\Lambda = \{\Lambda_{lj}\} = (Z^TZ)^{-1}Z^TX = VSV^T \tag{6}$$

where $l = 1, \dots, p, j = 1, \dots, p$. Because the variables in Z are orthonormal, $\Lambda_{lj}{}^2$ can be explained as the proportion of variance in x_j accounted for by Z_l.

Finally, the relative weight of *j*th variable in X is then defined as:

$$D_j = \sum_{l=1}^{p} \Lambda_{lj}{}^2 \hat{\beta}_{Z_l}{}^2 \tag{7}$$

In matrix notation, the relative weight of all variables in X can be obtained by:

$$D = (\Lambda \odot \Lambda)(\hat{\beta}_z \odot \hat{\beta}_z) = ((VSV^T) \odot (VSV^T))((VU^TY) \odot (VU^TY)) \tag{8}$$

where $\odot$ denote the element-wise multiplication. It can be shown that the sum of the relative weights is also equal to model's R^2 [2].

Relative Weight Analysis Under "Singular X_{n*p} With $n \geq p$" Condition

In the above section, we have reviewed relative weight analysis mathematically in regular condition, which the design matrix X_{n*p} is full column rank. However, in some situations, the complex manufacturing processes may have factors that are interrelated according to physical laws resulting in a design matrix without full column rank. In this section, we will extend the relative weight analysis to this kind of "singular $X_{n*p}$with $n \geq p$" condition. Noting that, in equation (2), $Z = XT$ indicates the feasible region of equation (2) is "orthogonal vectors in n dimension space which can be expressed as a linear combination of p columns of X". When X is not full column rank, however, it is impossible to obtain p orthogonal vectors through linear combination of columns of X. Thus, equation (2) can't be solved. To address this problem, we generalize the feasible region to "all orthogonal vectors in n dimension space". Then the best-fitting orthogonal transformation can be formulated as:

$$\begin{aligned} &\text{Min: } \operatorname{trace}(X-Z)^T(X-Z) \\ &\text{s.t.} \quad Z^TZ = I_p. \end{aligned} \tag{9}$$

Proposition 1:

Under "singular $X_{n*p}$with $n \geq p$" condition, the solution of optimization problem (9) exists but is not unique, and these solutions are:

$$Z = U^r V^{r^T} + Z_{null} \tag{10}$$

where r is the rank of X, U^r is n-by-r matrix consisting of first r left singular vectors, V^r is p-by-r matrix cosisting of first r right singular vectors, and Z_{null} is n-by-p matrix satisfying $Z_{null}V^r = 0_{n*r}$.

Note that Z_{null} is not unique which makes the solution of optimization problem (9) not unique. We will choose such a Z_{null} which makes the R^2 of model $Y = X\beta_X + \varepsilon$, denoted by ${R_X}^2$, equals to the R^2 of model $Y = Z\beta_z + \varepsilon$, denoted by ${R_Z}^2$. Noting that ${R_X}^2 = Y^TU^rU^{r^T}Y$ and ${R_Z}^2 = Y^TU^rU^{r^T}Y + Y^TZ_{null}{Z_{null}}^TY$, thus, ${R_X}^2 = {R_Z}^2$ implies ${Z_{null}}^TY = 0_{p*1}$. That is, we will choose such a Z_{null} that further satisfies ${Z_{null}}^TY = 0_{p*1}$. Then, the standardized regression coefficients of model $Y = Z\beta_z + \varepsilon$ can be written as

$$\hat{\beta}_z = \{\hat{\beta}_{z.l}\} = (Z^TZ)^{-1}Z^TY = V^rU^{r^T}Y + {Z_{null}}^TY = V^rU^{r^T}Y \tag{11}$$

Next, we regress X on Z to relate the orthogonal variables back to original variables

$$\Lambda = \{\Lambda_{lj}\} = (Z^TZ)^{-1}Z^TX = V^rS^rV^{r^T} \tag{12}$$

where S^r is r-by-r diagonal matrix with non-zero singular values of X on the diagonal.

With $\hat{\beta}_Z$ and Λ, the relative weights of all independent variables in X under "singular $X_{n*p}$with $n \geq p$" condition are then obtained by

$$D = (\Lambda \odot \Lambda)(\hat{\beta}_z \odot \hat{\beta}_z) = ((V^rS^rV^{r^T}) \odot (V^rS^rV^{r^T}))((V^rU^{r^T}Y) \odot (V^rU^{r^T}Y)) \tag{13}$$

Proposition 2:

The sum of relative weights under "singular X_{n*p} with $n \geq p$" condition is again equal to model's R^2.

$$\sum_{j=1}^{p} D_j = {R_X}^2 = {R_Z}^2 = Y^TU^rU^{r^T}Y \tag{14}$$

Relative Weight Analysis Under "$n < p$" Condition

In this section, we will extend the relative weight analysis to "$n < p$" condition. This will be done by converting the "$n < p$" condition into the "singular $X_{n*p}$with $n \geq p$"

condition with an expanded sample size. More specifically, we duplicate n observations of each variable in X m times to make the new sample size $mn \geq p$. With this procedure and normalization, we construct two new normalized data matrices denoted by X^*_{mn*p} and Y^*_{mn*1} as follows:

$$X^*_{mn*p} = \frac{1}{\sqrt{m}}\begin{pmatrix} X \\ \vdots \\ X \end{pmatrix}, Y^*_{mn*1} = \frac{1}{\sqrt{m}}\begin{pmatrix} Y \\ \vdots \\ Y \end{pmatrix} \tag{15}$$

The rank of X^*_{mn*p} is equal to the rank of X, denoted by r. It can be shown that ${R_X}^2$ equals to ${R_{X^*}}^2$, $\rho(x_j, Y)$ equals to $\rho(x_j^*, Y^*)$, $\rho(x_j, x_k)$ equals to $\rho(x_j^*, x_k^*)$, where ${R_X}^2$ is the R^2 of model $Y = X\beta_X + \varepsilon$, ${R_{X^*}}^2$ is the R^2 of model $Y^* = X^*\beta_{X^*} + \varepsilon^*$, and $\rho(\cdot)$ denotes the zero-order correlation. In other words, the correlation structures of X and Y are equivalent to the correlation structures of X^* and Y^*. Therefore, we claim that the relative weights of X is equivalent to that of X^*. Furthermore, with the following proposition, we find that all information of these relative weights is coming from the SVD of the original variables X and has no relationship with m.

Proposition 3:

$$D = D^* = ((V_x{}^r S_x{}^r V_x{}^{rT}) \odot (V_x{}^r S_x{}^r V_x{}^{rT}))((V_x{}^r U_x{}^{rT} Y) \odot (V_x{}^r U_x{}^{rT} Y)) \tag{16}$$

where *j*th element of D is the relative weight of *j*th variable in X.

Proposition 2 gives the fact that the sum of the relative weights of variables in X^* is equal to ${R_{X^*}}^2$ in that X^* belongs to the "singular X_{n*p} with $n \geq p$" condition. Thus, the sum of the relative weights under "$n < p$" condition is again equal to model's R^2 since relative weights of independent variables in X are equivalent to that of X^* and ${R_X}^2 = {R_{X^*}}^2$.

APPLICATION TO SEMICONDUCTOR YIELD LEARNING

In this section, a real semiconductor yield case with data provided by a local semiconductor foundry fab is used to validate the effectiveness of our methods in analysis of critical factors. There are nine inline measurement items of interest denoted by l_1~l_9, that may have an influence on an ET parameter which is the measurement of the current. The geometric relationship of the nine inline measurement items is represented in Figure 1.

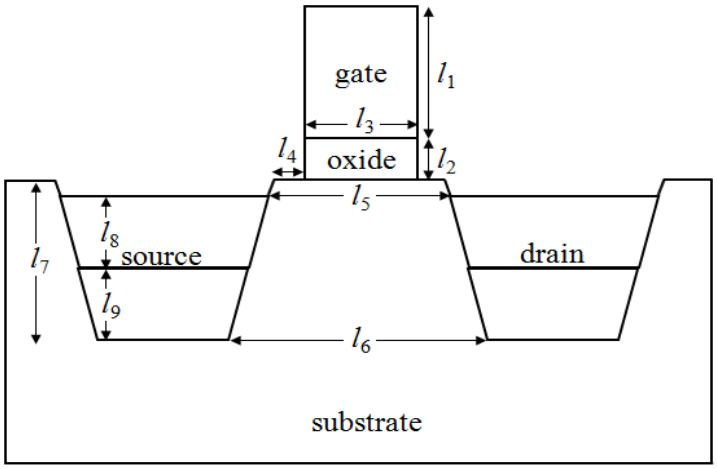


Figure 1: Geometric Representation of Inline Metrology Items l_1~l_9

As can be seen, there are two correlated sets. Items l_3, l_4, l_5 and l_6 are correlated among one another while items l_7, l_8 and l_9 forms another correlation set. The sample size is equal to 8, which is smaller than the number of independent variables. Table 1 shows the comparison of the independent variables in the order of their relative weights, the sequence of forward stepwise regression moves, the Lasso moves, and the least angle regression (LARS) moves. Relative weight analysis shows that the second set l_7~l_9 is the critical set compared to the first set l_3~l_6. In addition, l_1 and l_2 also have a critical impact on the ET parameter. The above results are consistent with both Lasso and LARS results where the second set l_7~l_9 are chosen before the variables in the first set l_3~l_6. However, the forward stepwise regression selects l_4 before l_8, which is not correct according to the physical laws.

Table 1: Comparison of Relative Weights, Forward Stepwise Regression, Lasso, and LARS

Methods	**Descending order of the variable selection moves**								
Relative Weights	l_9	l_8	l_2	l_1	l_7	l_3	l_5	l_6	l_4
Forward Stepwise Regression Moves	l_9	l_7	l_2	l_4	l_8	l_6	l_1	l_3	l_5
Lasso Moves	l_9	l_7	l_1	l_2	l_8	l_4	l_6	l_3	l_5
LARS Moves	l_9	l_7	l_1	l_2	l_8	l_4	l_6	l_3	l_5

CONCLUSIONS

In this research work, we have reviewed the history of calculating relative importance and extended the relative weight analysis to two low-rank conditions. Study on a real case has also demonstrated how the methodology can be used to discover the critical causes for semiconductor yield learning. The results obtained from the proposed method are consistent with the results of Lasso and LARS. As compared to classical forward stepwise regression, our method not only detects the correct importance order of variables but also has a more straightforward computation. Finally, we believe the proposed method can be further applied to the research area of variable selection especially for situations with the presence of multicollinearity.

REFERENCES

1. D.V. Budescu, “Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression”, *Psychological Bulletin*, 114, pp. 542-551, 1993.

2. J.W. Johnson, “A heuristic method for estimating the relative weight of predictor variables in multiple regression”, *Multivariate Behavioral Research*, 35, pp. 1-19, 2000.

3. J.M. Lebreton, R.E. Ployhart, R.T. Ladd, “A Monte Carlo comparison of relative importance methodologies”, *Organizational Research Methods*, 7, pp. 258-282, 2004.

4. R.M. Johnson, “The minimal transformation to orthonormality”, *Psychomatrika*, 31, pp. 61-66,1966.

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

Twitter Social Networking for Recommendation

Réda Soltani, Latifa Baba-Hamed
University of Oran 1 Ahmed Benbella
BP 1524 El M'Naouer, Oran, Algeria
reda.soltani@hotmail.fr, lbabahamed@yahoo.fr

Asma Djatou
University of Oran 1 Ahmed Benbella
BP 1524 El M'Naouer, Oran, Algeria
Asma.djatou@yahoo.fr

ABSTRACT

In this paper, we propose a recommendation system using data from social networks. Specifically, it recommends a list of experts, extracted from the social network Twitter, to a user who is interested in a specific topic such as sport, music, politic, education, etc. Our aim is to simplify access to the desired information for users of social networks.

Keywords: Recommendation, Twitter, Extraction, Classification-algorithm, Followers, Expert.

1 INTRODUCTION

The data coming from the social networks represent a wealth of information for a recommendation system. We can identify a new type of recommendation systems based on the presence of a community of users linked by social ties. In this paper, we want to make the recommendation using the extracted data from the social network Twitter. It is about the recommending experts in some fields (politics, sport, music, medicine, etc.) to a user making a request relating to one of these domains. Our aim is to simplify access to the desired information for users of social networks.

The rest of the paper is organized as follows. Section 2 presents the main works in the social recommendation. Section 3 is devoted to the architecture presentation of our system. Section 4 discusses the experiments conducted. Finally, section 5 concludes this article and presents some perspectives.

2 THE MAIN WORKS IN THE FIELD OF SOCIAL RECOMMENDATION

KeepUp [1] is a recommendation system for RSS[1] feeds articles. KeepUp is not a social network, it is a hybrid system that creates a social network implicitly and uses this social network to improve recommendations. We are only interested in its social aspect. In this system, each user can create Customized RSS channels (by adding tags to filter an RSS feed, for example). Channels created by a user are available to other users who can subscribe to them. They then receive all the new items from this channel.

FilmTrust [2] is a recommendation system where users are asked to explicitly quantify the confidence they have in their friends. Once Trust established the Trusted Relationship Inference Algorithm TidalTrust assigns a trust rating to peers that are not in the direct list of friends. This trust between nodes (individuals) of a trusted network makes it possible to calculate the weight granted by a user in the recommendation of a film.

Konjere and Dhawas [3] designed a system that generates privacy protection for friends' recommendation in online social networks, allowing two unknown people to establish trusted links in a small set of relationships of confidence. This system uses the social attributes of the users in social network to implement a multi-hop trust chain in the presence of each trust relationship 1-jump based on context where most of the trusted connections are formed and strengthened by shared social attributes.

Liben-Nowell and Kleinberg [4] formalize the problem of link prediction and develop an approaches based on proximity measurements of nodes in a network. In particular, they have experimented their method on a co-publication network of scientific papers. They propose two types of methods: the methods based on the neighborhood notion and the methods requiring the whole of the graph.

Lefeuvre and Cabanac [5] propose a method of analysis of social networks combining social data with thematic data related to published articles to make the recommendation of scientific collaboration. The idea is to define, in addition to a thematic function calculating a similarity between researchers, three other social functions: the inverse of the length of the shortest path in a graph of co-authors, the strength of the connectivity value enhancing capacity to reach a person via different intermediaries, and the number of joint conferences materializing a real meeting factor.

3 PROPOSED APPROACH

The aim of our approach is to recommend experts in certain fields such as politics, sport, music, medicine, etc., to a user making a request relating to one of these domains.

In fact, we have initially defined a list of categories (sport, politics, music, medicine, etc.) with specific keywords for each category. We have, in a second time, extracted a large number of Algerian tweets using rectangles method; it consists, to reject all the tweets not emanating from Algeria by the coordinates of these rectangles. Then, we take the text of the tweet and the keyword repository, in order to classify each tweet and know its category. The "keywords" method extracts all texts from the tweets one by one and compares them with the

[1] RSS (Rich Site Summary) is a format for delivering updates for website. RSS feeds are often used by news sites or blogs to present the titles of the latest information available online.

keywords for each category (sport, health, education, etc.). For each word found in the text of the tweet, we increment the counter of the related category. Once the words of the text are exhausted, we assign this tweet to the category whose counter is the highest.

At the end of the classification of the tweets by keywords, we obtain n sets of tweets corresponding to n different categories. Thus, we can extract the experts according to algorithm 1.

```
Algorithm 1 Experts extraction
C: set of categories
T: set of tweets
U: set of users who have issued tweets
S: threshold
Lc: list of experts to recommend for a given category c

Begin
Lc ← {}
For each uk in U do
NTuk ← 0      // Total number of tweets issued by the user uk
Nmax ← 0
For each ci in C do
NCuk ← 0      // Total number of tweets issued by the user uk
              // in category ci
   For each tj in T do
       If tj is emitted by uk
       then
       NCuk ← NCuk + 1
       End if
  End for
  Return NCuk
 NTuk ← NTuk + NCuk
if NCuk  > Nmax then
 Nmax ← NCuk
Cmax ← ci
Umax ← uk
End if
End for
if NTuk > S then
Lc ← Lc U {Umax}
Return (Umax, Cmax, Nmax/ NTuk)
End for
Return Lc ordered // in descending order according to the
                  // ratio Nmax/ NTuk
End
```

The algorithm 1 can be summarized by adopting, for each user U having an account in Twitter, the following steps:

- Calculate the number of tweets NC for each category
- Extract the category C in which the user U has the highest number of tweets

- Test if the user U has a total number of tweets N greater than a given threshold (10 for example to be able to consider U as an expert user in this category of themes).
- If yes, add U to the recommendation list Lc related to the category C

Once the expert extraction has been completed, the list Lc will be sorted in descending order according to the ratio F calculated by formula (1), where NC represents the number of tweets issued by the expert of a given category and N the total number of tweets issued by this expert. Figure 1 shows the general architecture of our system.

$$F = NC/N \qquad (1)$$

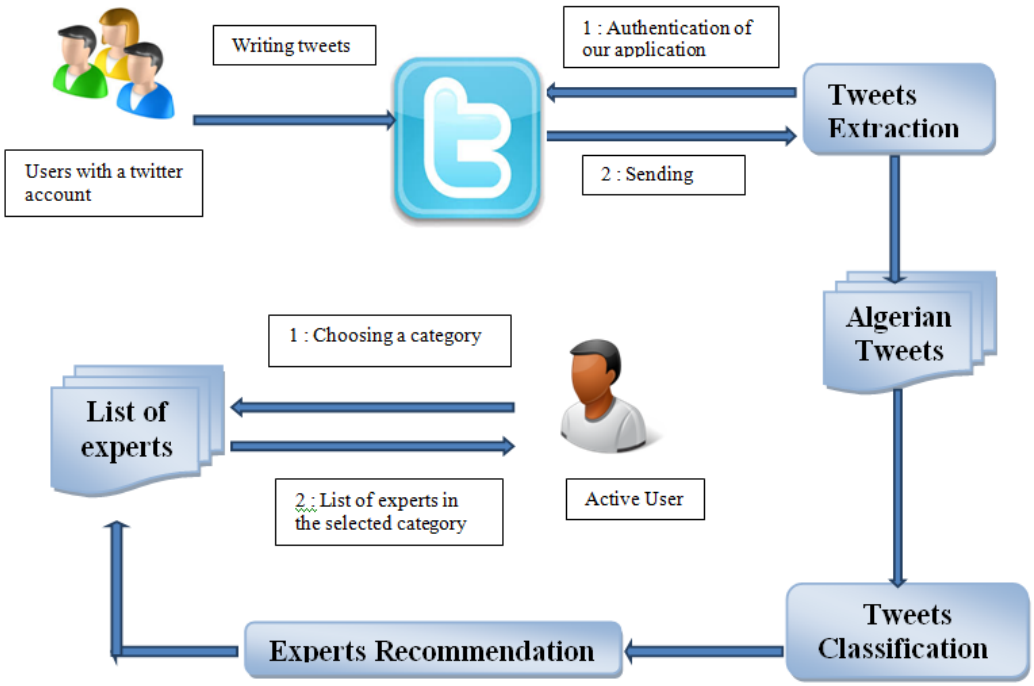


Figure 1: General architecture of the application

4 EXPERIMENTATION

In order to evaluate our method, we proposed to a population of people (students, friends) with age differences to give us a classification of 153 tweets. People's answers will be considered as a reference to evaluate the Keywords method.

After the classification of these 153 tweets and comparing them with the repository (emitted by people), we note that our method gave 65 correct answers on the 153 tweets.

Other classification methods have been developed before using two different algorithms, namely Bayesian networks and KNN (K nearest neighbors). These two algorithms are based on learning for each category. We executed them on the 153 tweets. We split this 153 tweets set into learning set and test set. Before proceeding to the evaluation given by each method, we note that:

- The evaluation of the keyword method is applied to all the153 tweets set.
- The methods evaluation (Bayesian networks and K Nearest Neighbor (KNN)) concerns one third of the tweets only (50 out of 153) for the test, the rest of the tweets (103) are used for training.
- Bayesian networks and KNN Methods are based on all the words that appear in the tweets of the training set.

4.1 EVALUATION : CALCULATION OF PRECISION, RECALL BY THE KEYWORD METHOD AND COMPARISON WITH BAYESIAN AND KNN METHODS

To evaluate our method, we used the two known functions in the domain of data retrieval: accuracy and recall based on people's answers.

The accuracy: it is the number of relevant tweets found relative to the total number of tweets proposed by the method. If it is high, it means that few answers, of our method, are false and can be considered as "precise". We calculate the accuracy with the formula 2.

$$\text{accuracy } i = \frac{\text{Number of tweets correctly assigned to category } i}{\text{Number of tweets assigned to category } i} \quad (2)$$

The recall: it is defined by the number of relevant tweets retrieved in relation to the total number of tweets proposed by people and by category. The formula 3 calculates the recall.

$$recall\ i = \frac{\text{Number of tweets correctly assigned to category } i}{\text{Number of tweets belonging to category } i} \quad (3)$$

F-measure: A popular measure that combines accuracy and recall and their weighting, called F-measure or F-score. The formula 4 calculates the F-measure.

$$F = \frac{2 .(\text{accuracy} .\text{recall})}{\text{accuracy} + \text{recall}} \quad (4)$$

A perfect classification system provides answers with accuracy and recall equal to 1 (the algorithm finds all the tweets in the correct category - recall - and does not make any errors as compared to what was expected - accuracy).

In reality, the classification algorithms are more or less precise, and more or less relevant. It is possible to obtain a very precise classification (for example an accuracy score of 0.99), but with poor performance (for example, with a recall of 0.10 which means that it has found only 10% of the possible answers). In the same way, an algorithm with a strong recall (for example 0.99 is almost all the relevant tweets), but the low accuracy (for example 0.10) provides as a response many classifications of erroneous tweets.

Thus, the value of a classifier is not reduced to a good score in accuracy or in a recall. The results of evaluations are summarized in Tables 1, 2 and 3, where:

NbReal: the number of tweets given by the people (considered as correctly classified)
NbMethod: the number of tweets given by one of the three methods (the keywords method, Bayesian Networks method and KNN method).

Table 1. Calculation of precisions, recalls and F1 for the keywords method

Category	nbReal	nbMethod	Intersection	Accuracy	Recall	F1
Sport	44	22	21	0.9545454545	0.4772727273	0.636363636
Politic	14	22	9	0.4090909091	0.6428571429	0.5
Technology	3	1	1	1	0.3333333333	0.5
Other	9	49	5	0.1020408163	0.5555555556	0.172413793
Culture	24	9	6	0.6666666667	0.25	0.3636363636
Social	37	27	7	0.2592592593	0.1891891892	0.21875
Religion	14	15	11	0.7333333333	0.7857142857	0.7586206897
Education	4	5	3	0.6	0.75	0.6666666667
Health	3	2	2	1	0.6666666667	0.8

Table 2. Calculation of recall, accuracy and F1 for the Bayesian Networks method

Category	nbReal	nbMethod	Intersection	Accuracy	Recall	F1
sport	14	32	14	0.4375	1	0.608695652
politic	5	0	0	0	0	0
technology	1	0	0	0	0	0
Other	3	0	0	0	0	0
Culture	8	1	0	0	0	0
Social	11	16	5	0.3125	0.4545454545	0.370370370
Religion	5	1	1	1	0.2	0.33333333
Education	2	0	0	0	0	0
Health	1	0	0	0	0	0

Table 3. Calculation of recall, accuracy and F1 for KNN Method

Category	nbReal	nbMethod	Intersection	Accuracy	Recall	F1
Sport	14	28	14	0.5	1	0.666666667
Politic	5	3	2	0.666667	0.4	0.5
Technology	1	0	0	0	0	0
Other	3	8	2	0.25	0.666666667	0.3636363636
Culture	8	4	2	0.5	0.25	0.333333333
Social	11	6	5	0.833333	0.45454545	0.588235291
Religion	5	1	1	0	0.2	0.333333333
Education	2	0	0	0	0	0
Health	1	0	0	0	0	0

4.2 RESULTS INTERPRETATION

We note in Table 1 that $precision_{religion}$ = 0.7333333333 and $recall_{religion}$ = 0.7857142857 means that the system is very precise and does not get too wrong. We explain this by the fact that keywords of the base are effective because the words that revolve around religion are known to all, we do not invent new words.

This gives us an $F\text{-measure}_{religion}$ = 0.7586206897, which means that our system is very efficient. Concerning the health category, the $precision_{health}$ = 1 which means that the system does not make any error with respect to what is expected. While $recall_{health}$ = 0.6666666667: the system finds 66% of the tweets in their correct category. We explain this by the fact that the key words of the database are effective, and that health is a precise domain, so we can not commit an errors classification. With $F\text{-measure}_{health}$ = 0.8, we deduce that our system is very efficient.

We observe in Table 2 that the $F\text{-measure}_{health}$ and $F\text{-measure}_{social}$ are very weak. This is due to the lack of learning and the quality of the tweets. $F\text{-measure}_{health}$ = 0, which means that there is a shortage in the number of learning tweets for the health category.

We note in Table 3 that the $F\text{-measure}_{religion}$ is weak for lack of agreement between the test and the learning. The $F\text{-measure}_{health}$ = 0, which means that there is a very small number of tweets of this category in learning. On the other hand, $F\text{-mesure}_{social}$ has a good score. That is, the tweets of learning and those of the results of people are more or less close.

We can't conclude that the keyword method is better than the other two classification methods but gives satisfactory results.

5 CONCLUSION

In this article, we presented our contribution which aims to recommend a list of experts extracted from the social network Twitter to a user who is interested in a specific topic such as sport, music, politics, education, etc. Our method combines the Community Detection and Behavior Analysis approach with the link prediction approach. Indeed, it is an approach that analyses the tweets to assign them to a specific theme. So by classifying all the tweets, and since we have limited the collection to a specific country, we can deduce the subjects addressed by the citizens of this country. This approach predicts, also links to users having accounts in Twitter. These links represent the experts of a given category, recommended to an active user.

The limitation of our method is the number of tweets, prospects are to be considered as:

- Use a benchmark larger than the one we have chosen
- Compare our method with one of the existing methods
- Increase the number of categories
- Combine two social approaches such as recommendation by choice of source with prediction of links
- Add semantics for categorization such as using ontology's, dictionaries, or expanding the keyword repository, and so on.
- Recommend expert users from other social networks such as: Facebook, or others.

6 REFERENCES

1. A. Webster, J. Vassileva, “The keepup recommender system. In Proceedings of the 2007 ACM conference on Recommender systems”, RecSys '07, pages 173-176, New York, USA, 2007.

2. J. Golbeck, J. Hendler, “Filmtrust : Movie recommendations using trust in web-based social networks”, In Proceedings of the IEEE Consumer Communications and Networking Conference, 2006

3. S. Konjere, V. N. Dhawas, “Online Social Network for Recommendation System using SVM. International Journal of Computer Applications (0975 – 8887) Volume 147 – No.14” 2016

4. D. Liben-Nowell, J. Kleinberg, “The link-prediction problem for social networks. Journal of the American Society for Information Science and Technology 58(7), 1019–1031”, 2007.

5. A. Lefeuvre, G. Cabanac. “Confronting the human perception of measures of similarity between members of an academic social network: enrichment of the theme by the social aspect. In MARAMI'10: Proceedings of the 1st Conference on Modeling and Analysis of Networks: Mathematical and Informatics Approaches ", 2010.

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

ATC work shift scheduling using multistart simulated annealing and regular expressions

F. Tello, A. Mateos, A. Jiménez-Martín, J.A. Fernández del Pozo
Departamento de Inteligencia Artificial, Universidad Politécnica de Madrid
Campus de Montegancedo S/N, Boadilla del Monte 28660, Spain
{faustino.tello, alfonso.mateos, antonio.jimenez, juan.fdezpozo.salamanca}@upm.es
web-page: http://www.dia.fi.upm.es/

ABSTRACT

In this paper, we propose a new approach for solving the air traffic controller (ATC) work shift scheduling problem, which minimizes the number of ATCs required to cover a given airspace sectoring, while satisfying a set of ATC labor conditions. This optimization problem belongs to the class of timetabling problems. The size and complexity of these combinatorial problems make them hard or even impossible to solve with exact methods.

In the proposed approach, initial feasible solutions are first built using a heuristic based on optimized templates, and then multistart simulated annealing is used to reach optimal solutions. In the search process, we use regular expressions to check the feasibility of the generated solutions. This provides high testing speed and modularity for a clear and maintainable implementation of the optimization model. Once the optimal ATC number is reached in one or more solutions, they are used as the initial solutions for a new optimization process aimed at balancing the ATC workloads.

The proposed approach is illustrated using a real example, and the optimal solution reached outperforms the reference solution, i.e. a real solution derived from the currently used tools based on templates. Indeed, one less ATC is needed to cover the airspace sectoring, and the ATC workloads are more balanced.

Keywords: Work Shift Scheduling Problem, Air Traffic Control, Simulated Annealing, Regular Expressions.

INTRODUCTION

The core of air traffic controller (ATC) activity is to facilitate the airspace and airport surface traffic flow under their responsibility, while avoiding collisions between aircrafts. To satisfy this essential safety constraint, they must detect and solve possible conflicts among trajectories.

The airspace is divided into sectors. These sectors are operated by two ATC working positions (executive and planner). All the sectors open at any one time must cover all the airspace. This is referred to as sectorization. The sectorization changes throughout the day depending on the aircraft traffic. A higher volume of air traffic involves opening more sectors and, consequently, more ATCs are necessary.

The sectorization required to handle the traffic estimations for a period can be designed beforehand. Therefore, a very important problem in air traffic control is to determine the minimum number of ATCs necessary to cover a sectorization structure for a given period, denoted as airspace sectoring, while satisfying certain strong constraints accounting for ATC labor conditions and real-time requirements for tactical decision level at the air control room.

This optimization problem belongs to the class of timetabling problems, which require the assignment of times and resources to events, considering sets of required and desirable constraints. The size and complexity (constraints) of these combinatorial problems make them hard or even impossible to solve with exact methods, like linear or constraint programming; or population-based metaheuristics, like genetic algorithms or particle swam optimization.

Different problem-solving approaches have been proposed in the literature to deal with timetabling problems [1]. Regarding the timetabling problem in the context of ATM, an overview of available results related to ATC shift scheduling is presented in [2]. There are already some software tools for generating ATC schedules [3], and their advantages and drawbacks have already been recognized [4]. Some are in-house tools, and details are not available for most of them. Therefore, it is necessary to build a decision support system to solving the ATC work shift scheduling problem, which minimizes the number of ATCs required to cover a given airspace sectoring, while satisfying a set of ATC labor conditions.

In this paper, we propose a novel methodology, wherein a heuristic is used to build initial feasible solutions and then multistart simulated annealing (SA) is used to reach an optimal solution. Regular expressions (Regex) [5] are used to check the feasibility of the visited solutions in the search process. The benefits of using Regex are high testing speed and modularity for a clear and maintainable implementation of the optimization model.

PROBLEM DESCRIPTION

One of the core tasks of ATCs is to avoid collisions among aircrafts. But there are two types of ATC. The executive ATC talks to the aircraft and gives instructions to the pilots to avoid conflict situations, whereas the planner ATC is responsible for anticipating possible conflicts and communicating with the executive ATC to solve the situation before it happens.

As cited before, the sectorization changes throughout the day depending on the aircraft traffic. The sectorizations needed to handle the traffic estimations for a time period (usually a day) can be designed beforehand. This is denoted as airspace sectoring. Thus, the airspace sectoring contains the sectors open over a twenty-four hour period.

Figure 1 shows an example of an airspace sectoring for Madrid Path1 in Spain. Each interval is associated with a configuration (1A,2A,4A…), where the integer value represents the number of open sectors and the letter refers to the sector configuration. Note that the 24-hour period has been divided into night (N)/morning (M)/afternoon (A) periods in Figure 1.

Besides, five different ATC shifts are considered: long morning (LMS) (5:40-14:00h.), morning (MS) (6:20-14:00h.), afternoon (AS) (14:00-21:20h.), long afternoon (LAS) (14:00-22:20h.) and night (NS) (21:20-6:20h.).

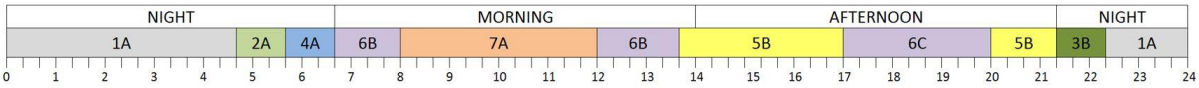


Figure 1: Madrid Path1 airspace sectoring

Moreover, a number of constraints accounting for ATC labor conditions have to be taken into account. In Spain, all these conditions were compiled and published in the Official State gazette (Boletín Oficial del Estado, BOE), Royal Decree 1001/2010, and Law 9/2010, regulating the provision of air traffic services. ATC labor conditions are as follows:

1. ATCs can cover a position (as executive or planner) in only one sector at any time.
2. Sectors must be covered by two ATCs (an executive and a planner ATC) during the time they are open.
3. ATCs must rest during 25% of the working shift in day shifts (MS, LMS, AS and LAS). A 33% resting time is established for the night shift (NS).
4. ATCs cannot work more than 2 consecutive hours.
5. Each rest period should last at least half an hour.
6. ATCs must remain in the same sector and position for at least half an hour (minimum time in a position).
7. Each ATC can work at most in two different sectors in the respective shift.
8. ATCs should occupy the planner (executive) position approximately 60% (40%) of the time.
9. ATCs cannot work more than 8 hours a day.
10. ATCs are assigned to one shift and cannot work in any other shift.
11. A sector change is not allowed without resting unless there is an emergency. ATCs can change to an affine (sector s1 and s2 are affine if $s1 \cap s2 \neq \emptyset$) sector without resting, but are required to work in the new sector for at least 15 minutes. Otherwise, ATCs need to rest to perform the sector change.

The objective of the problem is to determine the minimum number of ATCs necessary to cover a given airspace sectoring while satisfying the above ATC labor conditions. Moreover, it would be interesting to balance the ATC workload, avoiding big differences among ATCs. It was necessary to solve an instance of the problem at most in 15 minutes.

PROBLEM SOLVING METHODOLOGY

The proposed methodology is based on three elements, a heuristic to build initial feasible solutions, the use of Regex to check the feasibility of the solutions and a multistart SA to achieve optimal solutions.

Before describing the above elements in detail, let us first examine how the proposed methodology represents the solutions. We model the time discretely using a matrix containing

288 columns, i.e., the number of time slots under consideration. We consider time slots of 5 minutes since 5 is the greatest common divisor of 15 (11th ATC labor condition) and 20 (minimum period open a sector). Each row is associated with an ATC. The number of rows is established when applying the heuristic to build an initial solution. Note that initial solutions may have different numbers of ATCs.

Each element of the matrix (*i, j*) represents the state of the ATC *i* in the time slot *j*. We use the value 1 to represent a resting ATC, uppercase letters [A-Z] to point out that the ATC is working as an executive ATC in the corresponding sector, whereas lowercase letters [a-z] are used for planner positions. Finally, value 0 represents that the ATC is out of shift, i.e., neither working nor resting.

Colors are used to represent sectors when displaying solutions, see Figure 2. Resting periods and out of shift periods are in blue and white, respectively. The airspace sectoring in Figure 3 (11A, 13F, 11A) accounts for six hours rather than the whole day, and it is covered by 34 ATMs.

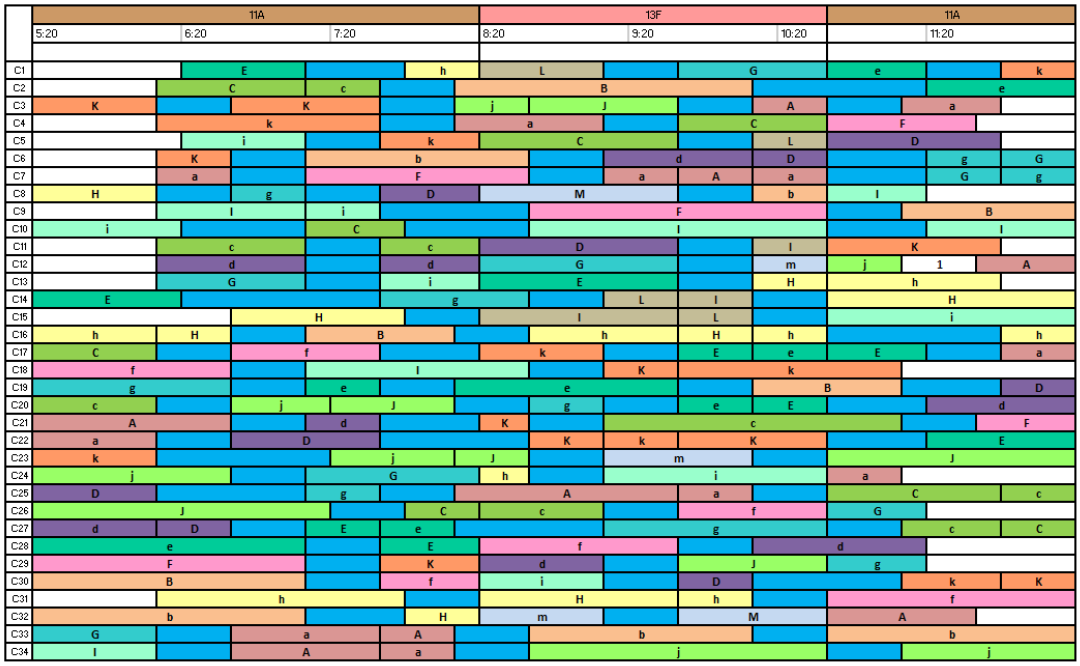


Figure 2: Solution representation

A heuristic for the construction of initial feasible solutions

We propose a heuristic to build a set of initial solutions with different resting periods. They will be used afterwards in the multistart simulated annealing to reach an optimum solution. The proposed heuristic is based on the use of an optimized template, see Figure 3, with three ATCs covering a sector during 96 time slots (eight hours).

In the template, the duration of the working periods is always the same and are twice as long as the resting periods. Moreover, if different resting period durations are used, the heuristic will build different initial solutions, albeit with a similar structure. As the labor conditions establish that each resting period should last at least half an hour (six slots), and the minimum and maximum working periods are six and 24 slots, respectively, the duration

of resting period must be between six and 12 slots. Thus, the heuristic build seven different initial solutions.

Note that some of the solutions output by the heuristic may not be feasible and, consequently, they are discarded as starting solutions for simulated annealing.

ATC 1	Working	Resting	Working	Working	Resting	Working	Working	…
ATC 2	Resting	Working	Working	Resting	Working	Working	Resting	…
ATC 3	Working	Working	Resting	Working	Working	Resting	Working	…

Figure 3: Template

Regular expressions

A regular expression (Regex) [5] is a special text string for describing a search pattern in texts. Regexs are composed of metacharacters \, (), [] and {}; characters classes like \A (start of string), \s (white space) or \d (digit), and quantiers * (0 or more), + (1 or more) and ? (0 or 1). For instance, the pattern "car(s)?" matches the words "car" and "cars", and (\w+@[a-zA- Z]+\.[a-zA-Z]{2,6}) could be used to match email addresses in a text, like Rose7@gmail.com.

In our ATC work shift scheduling problem, Regexs are used to check the feasibility of the solutions, i.e. the constraints representing the ATC labour conditions. For example, the labour condition that ATCs cannot work more than two consecutive hours can be verified using some patterns, i.e.: ^0*([a-zA-Z]{1,24}1{6,}[a-zA-Z]{1,24}1{6,}[a-zA-Z]{1,24})(0*$)

This pattern accepts strings between 1 to 24 (lower and uppercase) letters (5 minutes × 24 slots = 2 hours) following at least six “1” characters.

Regexs are also used to check that ATCs work in only one shift, ATCs do not work more than 24 consecutive slots (two hours), ATCs work at least six consecutive slots in the same position, no position change is performed without resting, and all opened sectors are covered by two ATCs (executive and planner).

Other labor conditions, including constraints regarding the total resting time, affine sectors treatment and the maximum number of work positions for each ATC cannot be verified using Regexs. Their verification has been implemented in the problem-solving algorithm.

Simulated annealing

SA [6, 7] is a trajectory-based metaheuristic which is named for and inspired by annealing in metallurgy. The basic idea of SA is as follows. An initial feasible solution is randomly generated. Then, in each iteration, a new solution is randomly generated from the neighborhood, of the solution considered in that iteration. If the new solution is better than the current one, then the algorithm moves to that solution. Otherwise, there is some probability of it moving to a worse solution. The acceptance of worse solutions makes for a broader search for the optimal solution and avoids trapping in local optima in early iterations.

The search is initially very diversified, since practically all moves are allowed. As the temperature drops, the probability of accepting a worse moves decreases, and only better moves will be accepted when it is zero.

The fitness function that we consider is maximizing $(h_1 + 2h_2 + \ldots + ch_c)/c^2$, where c is the number of ATCs and h_i is the number of slots that the i-th ATC is working. The term $1/c^2$ implies that when the number of ATCs decreases the fitness value is greatly improved. For a given number of ATCs, fitness values are higher for those solutions with a high workload for ATCs with higher indexes (in the last rows of the solution codification).

Thus, if we take into account the considered fitness function and the row reorganization of the visited solutions, we find that the search process tends to reallocate working periods from ATCs in the first rows to ATCs in the last rows (with the highest workload). In this way, we increase the resting periods among the first ATCs and tend to decrease the number of ATCs needed to cover the airspace sectoring.

The process for selecting a solution in the neighborhood of a given solution is as follows:

1. An ATC is selected at random. The aim is to reallocate some of this ATC's workload to another ATC.
2. A working period of the chosen ATC is selected at random. If it is composed of more than 12 slots, only the first 12 slots will be considered for reallocation, otherwise the whole working period is used.
3. We randomly select a second ATC and check if the whole working period for reallocation can be assigned to that ATC. If this is not possible, then we try with another ATC and so on. To do this, we check that the ATC under consideration is resting in the slots corresponding to the working period to be reallocated.
   If it is not possible to reallocate the considered working period to any ATC, then we reduce it by one slot (the last one) and repeat the process.
   We repeat the process until the working period is reallocated or is reduced to two slots, in which case, we go back to step 1.

ILLUSTRATIVE EXAMPLE

We consider one airspace sectoring to illustrate the proposed methodology. There are 11 open sectors from 5:20 to 8:20, 13 open sectors from 8:20 to 10:40 and 11 from 10:40 to 11:20, see Figure 2.

First, we run the heuristic to build the initial feasible solutions. Seven initial feasible solutions are obtained using different templates. Then, multistart SA is carried out with the following parameter values: we run the algorithm proposed in [8] with different initial temperatures for adjustment using an acceptance ratio 0.9, leading to T_0=0.75. The number of iterations during which the temperature does not change is L = 500 and we established α = 0.9. The search stops when the fitness of the best solution does not improve by at least 0.05% during 1750 iterations.

The multistart SA reached seven solutions with the following number of ATCs {35, 34, 34, 35, 34, 34, 34}. To do this, we used a PC Intel i5-3230M CPU on 2.60GHz with 8GB of RAM running on Windows 10. It took 0.17 minutes to build the initial solutions and 3.5; 4; 9.6; 4.3; 13.2; 3.7 and 4.4 minutes, respectively, to reach the seven optimal solutions.

Note that the reference solution includes 35 ATCs, i.e. the solution yielded by the tools available before the proposed approach. The average ATC workload and the standard deviation are 280 minutes and 43.72 for the reference solution, respectively.

The optimal solution outperforms the benchmark since 34 ATCs are necessary to cover the airspace sectoring, with an average ATC workload of 288.23 minutes and a standard deviation 64.28. Looking at the standard deviation, however, we realized that the ATC workload is more balanced in the reference solution.

Using the optimal solution as the initial solution in the new optimization problem aimed at balancing the 34 ATC workloads, we obtain the solution shown in Figure 2, with a lower standard deviation, 16.17.

CONCLUSION

We have proposed a new approach to solving the ATC work shift scheduling problem that minimizes the number of ATCs necessary to cover a given airspace sectoring while satisfying a set of ATC labor conditions according to Spanish regulations and real-time requirements. The approach consists of three elements: a heuristic to build initial feasible solutions, the use of Regex to check rapidly the feasibility of the solutions, and multistart SA to reach optimal solutions.

Once the optimal ATC number is reached, it is used as the initial solution of a new optimization process aimed at balancing the ATC workloads. The proposed approach has been illustrated using a real example, and the optimal solution reached outperforms the reference solution by one ATC and with more balanced workloads.

In this work, we have used simulated annealing for solving the corresponding optimization. However, we are currently implementing variable neighborhood search (VNS) and ant colony optimization (ACO) in combination with Regexs with the aim of comparing their performance.

Acknowledgement. The research reported in this paper was supported by Spanish Ministry of Economy and Competitiveness project MTM2014-56949-C3-2-R

REFERENCES

1. J. Telhada, “Alternative MIP Formulations for an Integrated Shift Scheduling and Task Assignment Problem”, Discrete Applied Mathematics 164, 328-343, 2014.

2. M. Arnving, B. Beermann, B. Koper, M. Maziul, U. Mellett, C. Niesing, J. Vogt, “Managing shiftwork in European ATM”, Literature Review, European Organization for the Safety of Air Navigation, 2006.

3. EUROCONTROL, “Shiftwork Practices Study - ATM and Related Industries”, DAP / SAF-2006/56 Brussels: EUROCONTROL, 2006.

4. EATCHIP Human Resources Team, “ATM Manpower Planning in Practice: Introduction to a Qualitative and Quantitative Staffing Methodology”, HUM.ET1. ST02.2000-REP-01 Brussels: EUROCONTROL, 1998.

5. J.E. Friedl, “Mastering Regular Expressions”, O'Reilly Media, Sebastopol, CA., USA 1997.

6. S. Kirkpatrick, C.D. Gelatt., M.P. Vecchi, “Optimization by Simulated Annealing”, Science 220, 671-680, 1983.

7. V. Cerny, “Thermodynamical Approach to the Traveling Salesman Problem: An Efficient Simulation Algorithm”, Journal of Optimization Theory and Applications 45, 41-51, 1985.

8. W. Ben-Ameur, “Computing the Initial Temperature of Simulated Annealing”. Computational Optimization and Applications 29, 369-385, 2004.

ICDSST 2017 PROCEEDINGS
THE EWG-DSS 2017 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
DATA, INFORMATION AND KNOWLEDGE VISUALISATION IN DECISION SUPPORT SYSTEMS
Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot, (editors)
Namur, Belgium, 29-31 May 2017

Flood risk assessment; A MCDM approach

Morteza Yazdani, Pascale Zarate
University of Toulouse, IRIT
2 Rue du Doyen-Gabriel-Marty, 31042 Cedex 9, Toulouse, France
morteza.yazdani@irit.fr, pascale.zarate@irit.fr

ABSTRACT

Uncertainty and risk are caused by million variables and either they will happen due to complex nature of the universe. This short communication investigates on effects of food risk variables (drives) on the agricultural section. Those drivers consider climate change, urban issues, socio-economic factors, and other measures. As the evaluation of agricultural production system based on multi variables is tough and usually confronts with conflicts, therefore, this article provides a framework to propose a structure with aid of multiple criteria decision making (MCDM) and expert attitudes. By an empirical example, eight decision factors and six alternatives compose a decision matrix and multi-objective optimization method (MOORA) delivers the optimal solution. The findings of this paper can be a route for experts in this area to explore the further questions and strategies.

Keywords: multi objective optimization based on ratio analysis, flood risk drivers, multiple criteria decision making, decision support model, RUC-APS

INTRODUCTION

Decision Support Systems (DSS) is a well-known approach introduced to offer the users the possibility of comparing options and computerize management decisions making activities through using information system technology. This concept has been integrated into many decision making applications involving multiple criteria decision-making (MCDM) to enhance capability and reliability of decision modeling [1] [2].

Risk and uncertainty are inevitable substances in agriculture and are appeared in terms of climate changes, soil erosion, water contamination, flood risk etc. These variables have the most significant impact on the ecosystem which should be addressed in agriculture decision making. The farmers and agriculture systems must realize and assess the risks and react to those risks through providing efficient strategies [3-5]. Recently a research project has been loaded called Risk and Uncertain Conditions for Agriculture Production Systems (RUC-APS) which mainly concentrates on finding the optimal solution for agricultural production systems through decision making and information technology assistance [6]. The project is supported by a scientific section of European Commission and guarantees to model decision support systems in order to deal with a sustainable agriculture supply chain once risk and uncertainty exist. This is a call of H2020-MSCA-RISE-2015 with the proposal number of 691249 in economic sciences panel. The project headed to improve agricultural conditions through

modeling well-structured decision process. Therefore, based on the RUC-APS objectives (**Figure 1**), to deal with a sustainable agriculture production system and to overcome those risks, establishing policy in land use management can be a challenge. In a sustainable agriculture, land use management and flood risk evaluation have essential affection. In this paper, we try to address a multi-criteria decision-making problem to build an initiative perspective in the evaluation of risk of disasters like the flood in agriculture production system.

Application of MCDM in agriculture decision making is growing quickly [7]. Adoption of integrated decision tools in terms of a decision support model to analyze decision problem, measure and formulate solutions can benefit users and decision experts to reach a robust land use strategy [8-10]. In this paper, it is possible to ask decision makers to present their comparison about risk factors importance and among decision alternatives in front of factors according to a predefined scale (see the **Table 1**). It will be reliable to count on Entropy technique, expert approach or other weighting process. Once decision makers report their opinions, then MOORA is able to produce desired results. Normally TOPSIS and VIKOR [11] tools are developed to find the optimal solution, however, in this study; we are going to experience a different taste of MCDM with the application of MOORA. In more recent works MOORA is going to be constituted by other well-known MCDM tools due to its simplicity and fast computation [12].

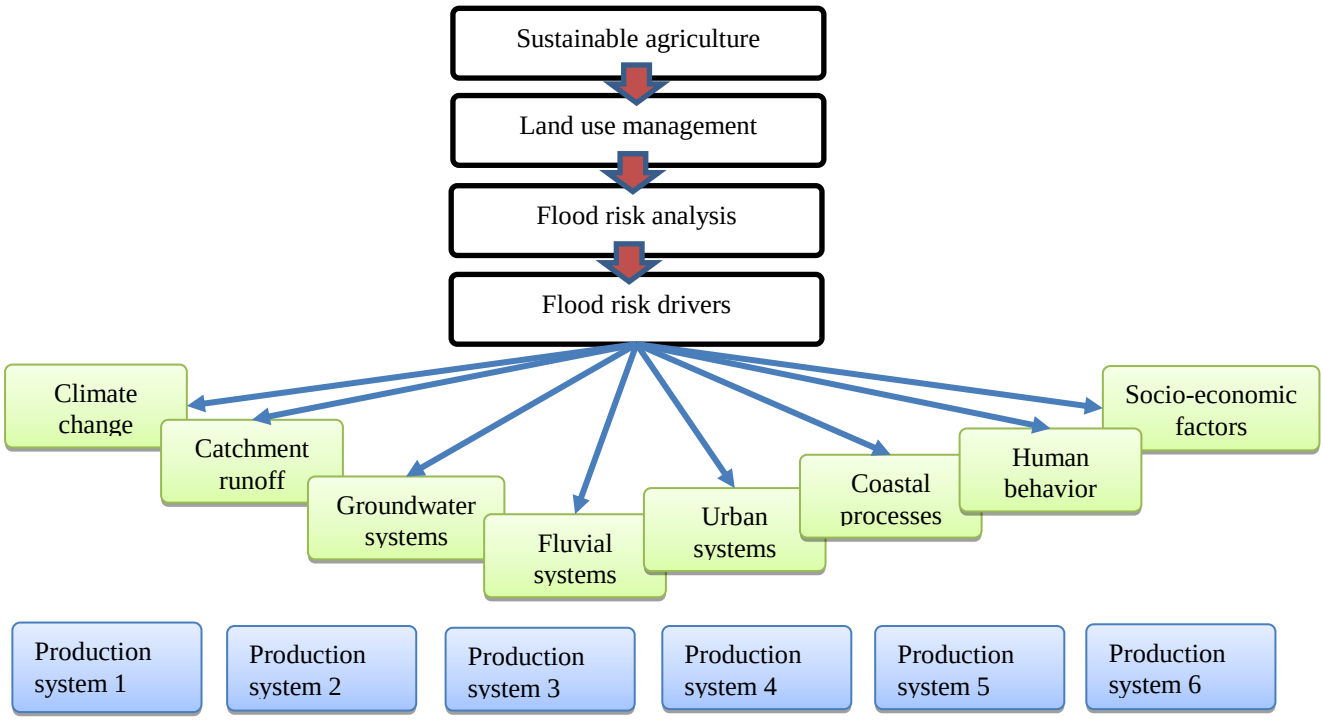


Figure 1: Analytical model for production system assessment

Proposed risk assessment model

This section undergoes to implement an evaluation frame by MOORA [13] method. To build decision making matrix firstly alternatives and criteria are detected, thereafter weights of the criteria and also rating of alternatives with respect to each decision criterion must be interpreted. For this paper, experts provide a rating scale to rate production systems

performance confronting flood risk drivers, and MOORA will compare and rank production systems. To evaluate flood risk affection, drivers (decision criteria) are defined as climate change (C_1), catchment runoff (C_2), groundwater systems (C_3), fluvial systems (C_4), urban systems (C_5), coastal processes (C_6), human behavior (C_7) and socio-economic factors (C_8) [5], six main agricultural production systems should be considered (alternatives) as it is observed in **Figure 1**. Among decision factors, C_1 is a cost factor because it is stated lower climate change lower risk for agricultural system and production. The decision making problem is this paper is solved by MOORA method. It has been claimed that considering interrelation between objectives and alternatives simultaneously, a cardinal approach and non-subjective dimensionless measures are the main characteristics of the MOORA. MOORA insists on two parts as; reference point approach and ratio system and is able to measure both non-benefit and benefit criteria in a process of selecting from a set of alternatives. Its process starts with identification of alternatives, detecting the most relevant criteria and determining the importance weights of criteria. The algorithm for MOORA is interpreted in this section [13]:

Step 0. Consider the following matrix is used to start the solution procedure: $X = x_{kj}$ (1)

It defines a decision matrix with k alternatives and j decision criteria

Step 1. Normalizing the decision matrix by: $r_{kj} = \dfrac{x_{kj}}{\sum_{k=1}^{t} x_{kj}^2} \quad k = 1,2,...t$ (2)

Step 2. Determining the weighted normalized matrix: $v_{kj} = r_{kj} \cdot w_j$ (3)

Step 3. Computing the overall rating of benefit and cost criteria for each alternative. The overall ratings of the k-th alternative considering the beneficial and non-beneficial criteria are calculated implementing;

$$S_k^+ = \sum_{j \in J^{Max}} v_{kj} \ , \ S_k^- = \sum_{j \in J^{Min}} v_{kj} \quad (4), (5)$$

where J^{Max} is the index set of the set of beneficial criteria and J^{Min} is the index set of the set of non-beneficial criteria

Step 4. Evaluating the overall performance of each alternative; $S_k = S_k^+ - S_k^-$ (6)

Step 5. Ranking the alternatives. The S_k values form a cardinal scale that can be used to rank the alternatives: the higher the value of S_k, the more preferred is the k-th alternative.

Results

An empirical study was considered to evaluate the risk of floods on production systems. The experts of the project are asked to present their judgments by relevant factors to evaluate alternative production system using **Table 1** predefined scale. In each driver, some sub-drivers are realized. For *climate change*: precipitation, temperature, relative sea-level rise, waves, surges; for *Catchment runoff*: urbanization and rural land management; Groundwater flooding; for *Fluvial systems*: environmental regulation, river morphology and sediment supply, river vegetation and conveyance, urbanization and Intra-urban Runoff; for *urban system*: sewer conveyance, blockage and sedimentation, Impact of external flooding on intra-urban drainage systems; for *coastal process*: coastal morphology and sediment supply; for

human preference stakeholder behavior have been considered. Additionally regarding the *socio-economic* factors these items were possible: buildings and contents, urban impacts, infrastructure impacts, agricultural impacts, social impacts, and science and technology [2],[4-5]. All of these issues are regarded and measured in the evaluation process.

The preference and overall judgments of decision makers should be provided using numerical scale in **Table 1** (how each system is influenced by a criterion). **Table 2** indicates the preference rating of decision makers (initial evaluation). The weights of the criteria can be obtained utilizing Entropy method. The information about Shannon Entropy method and its algorithm can be found in [14]. The generated weights by pairwise comparison are as; W_1 =0.084, W_2 =0.147, W_3 =0.056, W_4 =0.276, W_5 =0.177, W_6 =0.054, W_7 =0.128, W_8 =0.078. As observed the importance of fluvial system is much higher than others and this shows treatment with this system can affect risk of floods. The weights are utilized as importance level of risk drivers for MOORA process in order to determine the final solution.

Table 1: Relevant verbal reference to rate alternatives

Scale to rate production systems	
1	Very low
3	low
5	moderate
7	high
9	very high
For decision makers who decide values between each category 2,4,6 and 8 can be considered	

Table 2: Initial decision matric

	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
A_1	3	4	5	2	6	3	7	5
A_2	5	2	5	2	1	4	4	5
A_3	7	6	6	4	3	3	4	4
A_4	6	3	7	6	3	5	3	3
A_5	3	3	7	6	5	6	3	3
A_6	4	7	3	1	4	5	2	2

Table 3: Weighted normalized matric and ranking of alternatives

	Weighted normalized matrix								S^+	S^-	S_k	Rank
	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8				
A_1	0.021	0.0529	0.0203	0.0561	0.1081	0.0149	0.0884	0.0414	0.38204	0.021002	0.362	2
A_2	0.035	0.0264	0.0203	0.0561	0.018	0.0199	0.0505	0.0414	0.23262	0.035003	0.1977	6
A_3	0.049	0.0793	0.0244	0.1122	0.0541	0.0149	0.0505	0.0331	0.36842	0.049004	0.3195	4
A_4	0.042	0.0396	0.0284	0.1683	0.0541	0.0249	0.0379	0.0248	0.37798	0.042004	0.336	3
A_5	0.021	0.0396	0.0284	0.1683	0.0901	0.0298	0.0379	0.0248	0.41899	0.021002	0.398	1
A_6	0.028	0.0925	0.0122	0.028	0.0721	0.0249	0.0253	0.0165	0.27146	0.028003	0.24346	5

In this step, we can solve MOORA decision problem by algorithm presented in previous section and formulas 5-10. Firstly normalized matrix is delivered; thereafter weights of decision factors must influence decision process. In the further action overall rating for benefit and cost criteria are measured and validated based on weighted normalized matric values (**Table 3**). Ultimately, ranking of production systems can be derived based on higher values of S_k which are depicted here (see the **Table 3**);

$A_5 > A_1 > A_4 > A_3 > A_6 > A_2$

The ranking order of the alternatives shows the optimal option for agricultural objectives based on risk drivers of flood. The benefit of the results is that users in this project can figure out a better view on different variables, their relationship and affection on main objective of the research. For example from the climate change issue, it is a deal to support decision and assessment process because it has impacts on soil structure, biogeochemical cycles, and hydrological processes.

CONCLUSIONS

MCDM aims to develop models and structures to offer a better understanding of the decision system, feasible solutions, interrelationship of factors and windows for further improvements. This is core contribution of the MCDM modeling. This paper originates by an assessment approach based on multiple criteria decision making methods. It is investigated to develop a sustainable agriculture management, considering flood risk and its management is a significant topic. Therefore, it will demand an evaluation system to study the influence of the flood risk drivers with respect to several alternative production systems. We tried to improve better perspective of drivers such as climate change and its impacts on the production system based on experts judgment. To head that goal, a decision making problem regarding multiple factors has been structured. The relative importance of risk drivers has been obtained by experts' interaction and preferences. Then a decision table was built to deal with the affection of those factors with agriculture production system. Finally, by a newborn multi objective decision making tool, the ranking list of alternative projects has been announced. The contribution of this work will be implemented in RUC-APS project to analyze alternatives projects and warn the stakeholders and partners about possible corrective reactions. We have shown how conflicting factors can come together to the judgments of decision makers and through an MCDM framework. This is an initial but potential study for realization the interaction and influence of different factors related to risk and danger of floods and another kind of disasters. The users and partners in RUC-APS projects can take advantage of this model, and then is easy to implement, extend or combine it with other research projects. Integrated multi attribute modeling [15-16] with aid of the engineering tool can enhance the reliability of the work and decrease complexity.

The configuration of a decision support system allows the managers and policy makers to make effective decisions. Moreover, a strong decision support system requires a comprehensive and understandable decision framework. In this paper we have provided a decision model which is able to give this chance to the experts and managers to confront with the risk of flood and natural disasters. The managers can consider the proposed MCDM algorithm as a primal perspective for further improvement and possible extensions.

REFERENCES

1. Ngai, E. W., & Wat, F. K. T. (2005). Fuzzy decision support system for risk analysis in e-commerce development. *Decision support systems*, *40*(2), 235-255
2. Levy, J. K. (2005). Multiple criteria decision making and decision support systems for flood risk management. *Stochastic Environmental Research and Risk Assessment*, *19*(6), 438-447.
3. Mertz, O., Mbow, C., Reenberg, A., & Diouf, A. (2009). Farmers' perceptions of climate change and agricultural adaptation strategies in rural Sahel. Environmental management, 43(5), 804-816.
4. Malczewski, J. (2004). GIS-based land-use suitability analysis: a critical overview. Progress in planning, 62(1), 3-65.
5. Wheater, H., & Evans, E. (2009). Land use, water management and future flood risk. Land Use Policy, 26, S251-S264.
6. RUC-APS, 2016; http://cordis.europa.eu/project/rcn/204775_en.html
7. Jozi, S. A., & Ebadzadeh, F. (2014). Application of multi-criteria decision-making in land evaluation of agricultural land use. Journal of the Indian Society of Remote Sensing, 42(2), 363-371.
8. Chung, E. S., & Lee, K. S. (2009). Prioritization of water management for sustainability using hydrologic simulation model and multi criteria decision making techniques. Journal of Environmental Management, 90(3), 1502-1511.
9. Collins, M. G., Steiner, F. R., & Rushman, M. J. (2001). Land-use suitability analysis in the United States: historical development and promising technological achievements. Environmental management, 28(5), 611-621.
10. Uricchio, V. F., Giordano, R., & Lopez, N. (2004). A fuzzy knowledge-based decision support system for groundwater pollution risk evaluation. Journal of environmental management, 73(3), 189-197.
11. Yazdani, M., & Payam, A. F. (2015). A comparative study on material selection of microelectromechanical systems electrostatic actuators using Ashby, VIKOR and TOPSIS. Materials & Design (1980-2015), 65, 328-334.
12. Brauers, W. K. M., & Zavadskas, E. K. (2008). MULTI-OBJECTIVE OPTIMIZATION IN LOCATION THEORY WITH A SIMULATION FOR A DEPARTMENT STORE. Transformation in Business & Economics, 7(3).
13. Brauers, W. K. M., & Zavadskas, E. K. (2006). The MOORA method and its application to privatization in a transition economy. Control and Cybernetics, 35(2), 445.
14. Shemshadi, A., Shirazi, H., Toreihi, M., & Tarokh, M. J. (2011). A fuzzy VIKOR method for supplier selection based on entropy measure for objective weighting. Expert Systems with Applications, 38(10), 12160-12167.
15. Ignatius, J., Rahman, A., Yazdani, M., Šaparauskas, J., & Haron, S. H. (2016). An integrated fuzzy ANP–QFD approach for green building assessment. Journal of Civil Engineering and Management, 22(4), 551-563.
16. Yazdani, M., Chatterjee, P., Zavadskas, E. K., & Zolfani, S. H. (2017). Integrated QFD-

MCDM framework for green supplier selection. Journal of Cleaner Production, 142, 3728-3740.

ICDSST 2017 Posters

ICDSST 2017 Data, Information and Knowledge Visualisation in Decision Support Systems

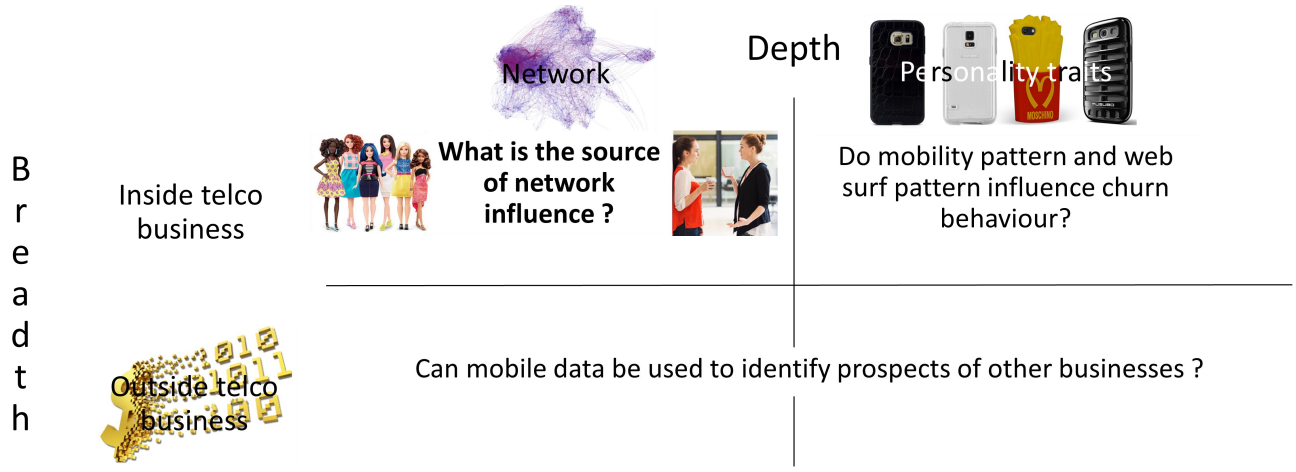
DISENTANGLING HOMOPHILY FROM SOCIAL INFLUENCE THROUGH UNDERLYING NETWORKS IN CUSTOMER RELATIONSHIP MANAGEMENT

Christian Colot*, Philippe Baecke**, Isabelle Linden*

****UNamur, rue de Bruxelles 61, 5000 Namur, Belgium***
***** Vlerick Business School, Reep 1, 9000 Gent, Belgium***
Christian.colot@unamur.be, philippe.baecke@vlerick,com, isabelle.linden@unamur.be
web-page: https://icdsst2017.wordpress.com/

Last decade has witnessed a growing interest in social network analysis (SNA) in the context of Customer Relationship Management (CRM). The emergence of device mediated social networks such as digital social networks is of special interest for marketers for its potential of using it to detect and to influence world of mouth linked to the company and/or to benefit from some homophily phenomenon i.e. people belonging to the same community may share some similarities such as the same attitude with respect to a specific offer.
Disentangling these two sources of convergent behaviors between a person and his/her peers is a hard task. It is however crucial to solve this issue to develop an effective CRM contact strategy. In this research, we propose to conduct this dissociation by comparing different networks among the same individuals. Three networks are investigated: spatial, communication and referral of a telephone company. These networks have different origins that are reflected in differing characteristics.
These different origins lead to the networks being representations of different underlying social forces. Whereas spatial networks are representations of homophily because similar people live close to each other, communication and referral networks resemble social contagion as they portray interactions among individuals and active recommendations. Spatial networks usually have many edges because every individual is connected to every other individual in the same neighborhood. Communication networks typically also have a large number of connections, although less than spatial networks. Referral networks have the least number of edges since people generally make less referrals than they have communication interactions. However, connections in referral networks typically are stronger than communication and spatial links because it requires the largest social investment.

POSITIONING OF RESEARCH QUESTION WITHIN CORRESPONDING RESEARCH PROJECT



OPEN ISSUES

- Research for best modelling techniques/metrics to measure the importance of each network
- Research for new telecom partners

A Collaborative Solution for IT Infrastructure Maintenance Based on Web Services and Mobiles Agents

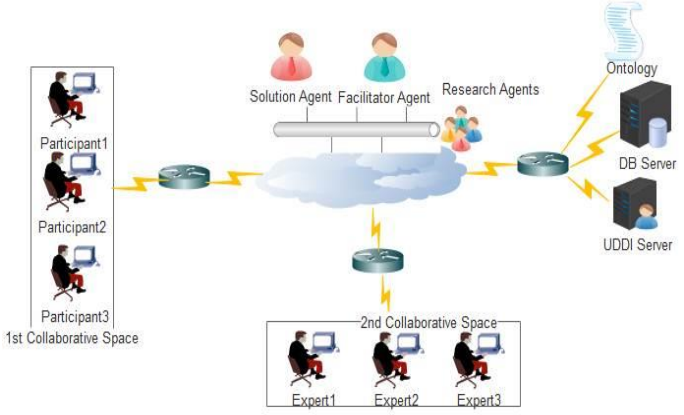
Elandaloussi Sidahmed, Taghezout Noria
Ahmed Ben Bella/Oran1 University
B.P 1524 M'Naouer, Oran 31000, Algeria
s-elandaloussi@hotmail.com, Taghezout.nora@gmail.com

Zaraté Pascale
IRIT /Toulouse University
2 rue du Doyen Gabriel Marty, 31042 Toulouse Cedex 9, France
Pascale.Zarate@irit.fr
web-page: https://www.irit.fr/~Pascale.Zarate

Abstract

Today, organizations have become more and more complex which has complicated the decision making process. However, with the technological advances and the evolution of collaboration technologies, industrial companies wish to benefit from it, in order to accelerate the interventions of the specialists of maintenance on the sites and to reduce the time inactivity of their equipment. As a result, this makes it possible to envisage keeping them at a distance without necessarily having locally qualified staff. Indeed, these situations require the rapid and effective intervention of experts, who are not basically available. Furthermore, in this work we propose a new service oriented approach which is essentially based on mobiles agents.

Global Architecture

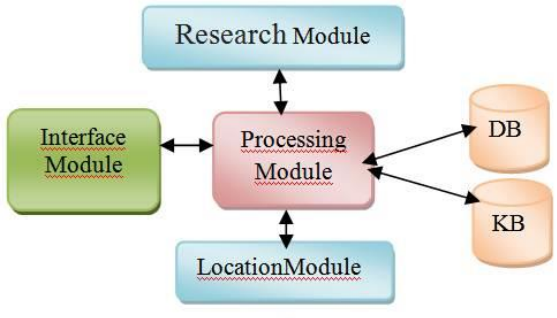


•Research Agents: These mobile agents are responsible for searching information (Web services) asynchronously.
•Solution Agent: This agent is responsible for seeking solutions to the problem that is posed.
• Facilitator Agent: The main task is to control the whole mechanism of decision-making.
•Database, Servers, Ontology of the domain, Directory UDDI.

Proposed Approach

•The development of a collaborative platform that facilitates a collective decision-making .
•Invoke and execute some WEB services depending on the problem nature.
•Integrates some mobiles agents in order to take benefit from the characteristics of mobile agents and to minimize response time.

Research Agent Structure



DB= Data Base ; KB= Knowledge Base ;

•Responsible for searching information (Web services) asynchronously.
•Move between the sites that are related to the company.
•Analyze the problem and to avoid the possibility of having a bad identification of the problem.
•Able to work more efficiently: rather than accessing this resource through a WAN.

Collaborative Spaces

•An internal collaborative space: Which consists of a set of participants by using a set of tools which are offered by the developed tools (Forum discussion, Shared calendar ....).
•A set of external decision-makers: represents the experts who offer solutions (with the recommendation of some web services).

A General View Of Our Collaborative Space

•In Our System, we use some tools as follows: Jade platform for multi-agent system development and Wamp server for databases implementation, Java Netbeans environment and Apache Tomcat for web pages deployment.

CONCLUSIONS

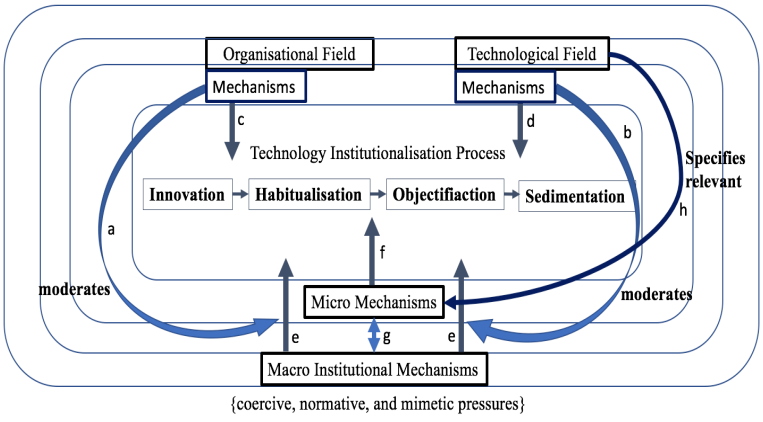
Our approach will enable us to reinforce the communication and cooperation of the various actors that are involved in the decision-making process and also allow them to consolidate their points of view by using multiple services which are developed and proposed by experts in the field. The integration of mobile agents increases the flexibility and reliability of the system and minimizes the response time by seeking the best service according to specific methods. As a result, it has allowed us to avoid the bad identification of the problem by the participants.

Information Technology Institutionalisation: A Multi-Level Process Model

Mathias Hetzel
Cass Business School, City, University of London
106 Bunhill Row, London, EC1Y 8TZ, UK
Mathias.hetzel.2@cass.city.ac.uk
www.cass.city.ac.uk

I develop a process model that addresses the call for more multi-level approaches to examine the technology institutionalisation process. The model incorporates micro, field, and macro-level mechanisms and the interaction effects between them. It centres on the concept of technological fields, which are structured around the production, use, definition, and control of a specific technology such as a Decision Support System (DSS). I propose to apply the notion of a technological field, representing the level between the society and an organisation, to address the neglected role of materiality in institutional processes and the inherent dualism between the social implications of technology use and the material aspects of technology design. The lack of focus on materiality in organisational studies is addressed by conceptualising technology as an institution and by emphasising the moderating role of the technological field in the institutionalisation process. The model developed in this paper describes how information technology (IT) becomes institutionalised and as such a taken-for-granted source for organisational decision-making. The model emphasises the importance of technology design for the institutionalisation process. The way a DSS visualises information significantly influences how likely it can become institutionalised. This study stresses the need to better understand the mechanisms that drive the institutionalisation of IT, in particular of systems such as DSS, that have a direct influence on decision-making processes. This knowledge contributes to designing tools that become an accepted and trusted information source.

THE MULTI-LEVEL TECHNOLOGY INSTITUTIONALISATION MODEL



a, b = field-level moderators
c, d = field-level direct mechanisms
e = macro-level mechanisms
f = micro-level mechanisms
g = micro/macro interaction effects
h = technological field specifies relevant micro mechanisms

THE TECHNOLOGY INSTITUTIONALISATION MODEL:

At the core of the model is the technology institutionalisation process, which has been adapted from Tolbert & Zucker (1999). The process starts with an innovative DSS that goes through three-stages consisting of habitualisation, objectification and sedimentation until full institutionalisation is achieved. Habitualisation refers to the formation of problem-solving behaviours and their association with specific stimuli. Objectification is the development of common social meanings related to these behaviours [8]. Sedimentation describes the processes through which actions obtain exteriority [8]. Exteriority relates to the level to which typifications are "experienced as possessing a reality of their own, a reality that confronts the individual as an external and coercive fact" [9: 58]. This process is influenced by macro, field (organisational and technological), as well as micro-level mechanisms, and the interaction effects between them. A DSS needs to adopt a certain stability and taken-for-grantedness to become institutionalised. An institutionalised DSS becomes part of the formal functioning of organisations, is embedded and aligned with the organisation's operations, is used for important tasks, is familiar, and its usage becomes intuitive. The DSS and users become deeply intertwined [10,11].

Selected Mechanisms Leading to Institutionalisation

Mechanism	Type	Description
Fields with impermeable boundaries	a	The strength of isomorphic forces is higher in organisational fields with impermeable boundaries than in more permeable fields [1]
Technology alternatives	b	A technological field (e.g. Financial Decision Support Systems) with relatively few alternatives experiences stronger isomorphic pressures for technology institutionalisation [1]
Regulatory agencies	c	Regulatory mechanisms specific to the organisational field (e.g. financial services industry) leading to organisational change [2]
Complementarity	d	A DSS that is compatible with a specific technological standard in the technological field is more likely to become institutionalised [3]
Design	d	The compatibility among technology components, achieved through a design reflecting a common standard [4]
Institutional Pressures	e	Coercive, normative, and mimetic pressures, e.g. a DSS is adopted and implemented with the desire of organisations to follow the actions of competitors (mimetic) [5]
Business alignment	f	The DSS was developed for a specific purpose, reflecting a particular business need [6]
Familiarity	f	The DSS replicates the values and beliefs of the organisation through the use of colours, logos, information structure and branding guidelines [6]
Macro/micro mechanisms interplay	g	Habitualised actions and behaviours at the micro-level are transformed into macro-level structure [7]

CONCLUSIONS & REMARKS:

This article argues that a better understanding of the technology institutionalisation process requires a multi-level perspective of micro, field, and macro factors. Existing research on institutionalisation processes is extended by suggesting a multi-field perspective. The technological field concept allows to identify field-specific mechanisms that contribute to the institutionalisation of technology. A major aim of this research has also been to encourage organisational scholars to a more active engagement with technology as well as to stress the significance for information systems scholars to better understand how IT can become an accepted and trusted information source.

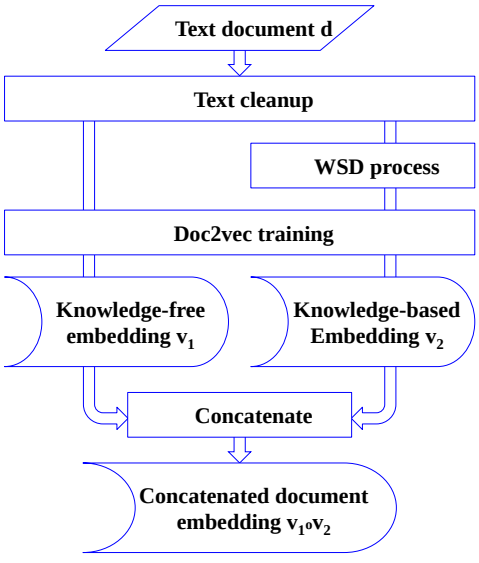
References are available on request

Concatenating Knowledge-based to Knowledge-free Document Embeddings for Text Classification

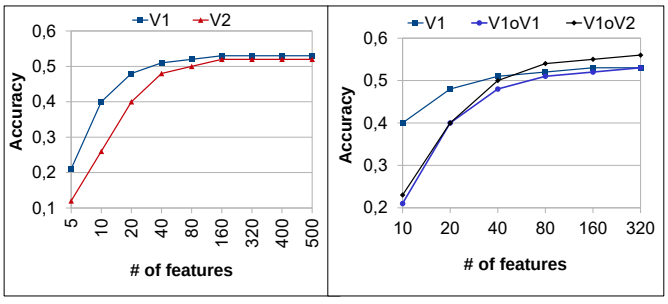
Hamid Machhour, Ismail Kassou
ENSIAS - UM5R
Madinat Al Irfane, BP 713, Agdal Rabat, Morocco
hamid.machhour@gmail.com, i.kassou@um5s.net.ma
web-page: http://ensias.um5.ac.ma

Document embedding has gained a lot of interests in text classification area. This paper investigates the popular neural document embedding method Paragraph Vector, publicly known as Doc2vec, as a source of evidence in document ranking. We focus on the effects of combining knowledge-based with knowledge-free document embeddings for text classification task. We concatenate these two representations so that the classification can be done more accurately. The results of our experiments show that this approach achieves better performances on a popular dataset.

STEPS TO BUILD DOCUMENT EMBEDDINGS



ACCURACIES OF LR CLASSIFIER ACCORDING TO DOCUMENTS EMBEDDINGS v_1, v_2 AND $v_1 \circ v_2$



BRIEF DESCRIPTION:

- **Text cleanup**: some primitive steps such as words filtering are needed to cleanup noises and reduce the vocabulary size.
- **Doc2vec training:** gives a weight matrix in which entries correspond to documents embeddings (document-vectors v_i)
  - *Knowledge-free document embedding* is obtained by training a Doc2vec model without using any external resources.
  - When additional preprocessing steps, using external resources, are invoked before Doc2vec training, we get the *knowledge-based document embedding*.
- **Word Sense Disambiguation (WSD) process** : senses of words are obtained by applying the original Lesk algorithm. Then, words are replaced in text by their senses.
  - In deep of WSD process, a Wordnet-based Text Tagger is used, and only nouns and adjectives are replaced by their senses in this approach.

After building document embeddings, a text classifier was trained and tested to expose the gain of concatenating embeddings v_1 and v_2 as $v_1 \circ v_2$.
For computations speed, we have used a simple linear regression (LR) classifier.

CONCLUSIONS & REMARKS:

- Accuracies of LR classifier based on vectors v_1 and $v_1 \circ v_1$ are used as baselines.
- Experiment shows slightly low accuracies when v_2 is used as representation of documents with respect to the baseline v_1
- Better accuracies are obtained with $v_1 \circ v_2$ representation, which can be up to 3% especially when the number of features is chosen between 100 and 500, the recommended range for document embedding size.
- Accuracy can be improved with better WSD algorithms.

In this work, we have investigated Paragraph Vector, publicly known as Doc2vec, to build a new document embedding by concatenating knowledge-free document embedding and knowledge-based document embedding. This approach was tested on a text classifier to expose the gain of concatenating these two representations. The conducted experiments show that concatenation gives more classification accuracies than using each representation separately.
Relevance of the presented approach is strongly related to the accuracy of the WSD process used to build knowledge-based document embedding. So, we anticipate test other WSD algorithms in our future works.

ICDSST 2017 Data, Information and Knowledge Visualisation in Decision Support Systems

Decision Support System for medication change of Parkinson's disease patients: a state-transition model

Biljana Mileva Boshkoska[1,2], Dragana Miljković [1], Anita Valmarska [1,5], Dimitris Gatsios [3], George RIigas [3], Spiros Konitsiotis [4], Kostas M. Tsiouris[3] , Dimitrios Fotiadis [3], Marko Bohanec [1]

[1]Jožef Stefan Institute, Ljubljana, Slovenia

[2]Faculty of information studies in Novo mesto, Slovenia

[3]Unit of Medical Technology and Intelligent Information Systems, University of Ioannina, Greece

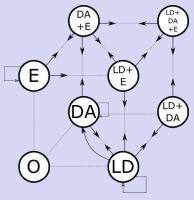
[4]Department of Neurology, Medical School, University of Ioannina, Ioannina, Greece

[5]Jožef Stefan International Postgraduate School, Ljubljana, Slovenia

[biljana.mileva, dragana.miljkovic, anita.valmarska]@ijs.si, dgatsios@cc.uoi.gr, george.a.rigas@gmail.com, skonitso@gmail.com, tsiourisk@biomed.ntua.gr, fotiadis@cc.uoi.gr, marko.bohanec@ijs.si

Parkinson's disease (PD) is a degenerative disorder of the central nervous system. Currently there is no cure for PD hence it requires a long-term, interdisciplinary disease management including typical medicament treatment with levodopa (LD), dopamine agonist (DA), and enzymes (E), such as MAO-B inhibitor. Due to the different combinations of motor and mental symptoms from which PD patients suffer, in addition to existing comorbidities, the interchange of medications and their combinations is patient-specific. In the frames of EU Horizon 2020 project PD_manager (http://www.parkinson-manager.eu/) we developed a decision support system for PD management which suggests *how* to change the medication treatment given patients' current state. The assessment is based on data that includes patients' motor symptoms (dyskinesia intensity, dyskinesia duration, offs duration), mental problems (impulsivity, cognition, hallucinations and paranoia), epidemiologic data (patient's age) and comorbidities (cardiovascular problems, hypertension and low blood pressure). The model is developed using DEX methodology, which integrates the qualitative multi-criteria decision modeling with rule-based expert systems. The model is composed of (1) a state-transition model that presents the medication change among levodopa, dopamine agonist, MAO-B inhibitors and their combinations, and (2) decision rules for triggering the changes, represented in terms of a qualitative DEX model.

A state-transition model for medication change (1a)



When suggesting medication change among levodopa (LD), dopamine agonist (DA), MAO-B inhibitors (E) and their combinations it is important to detect conditions that indicate a possible need to make a transition from some state to another. Thus, each arc in the state-transition model has an associated set of decision rules that determine whether the transition is desired or not. Symbol O represents the state where the patient does not take medications.

Structure and value scales of the "How" medication change DSS (2a)

Attribute	Scale
ChangeDAtoLD	yes; *no*
├*usingDA*	yes; *no*
├*MentalProblems*	yes; *no*
├*cardiovascular*	yes; no
└*low blood pressure*	yes; no
ChangeDAtoDA+LD	yes; *no*
├*usingDA*	yes; *no*
├*maxDA*	yes; no
└*MotorSymptoms*	yes; *no*
ChangeDAtoDA+MAOI	yes; *no*
├*usingDA*	yes; *no*
├*MotorSymptoms*	yes; *no*
└*cardiovascular*	yes; no
DecreaseDAdosage	yes; *no*
├*usingDA*	yes; *no*
├*MentalProblems*	yes; *no*
├*cardiovascular*	yes; no
└*low blood pressure*	yes; no
IncreaseDAdosage	yes; *no*
├*usingDA*	yes; *no*
├*maxDA*	yes; no
├*MotorSymptoms*	yes; *no*
├*offs duration*	yes; no
├*MentalProblems*	yes; *no*
├*cardiovascular*	yes; no
├*age*	lt65; 65-75; gt75
└*activity*	yes; no

ChangeLDtoLD+DA	yes; *no*
├*usingLD*	yes; *no*
├*MotorSymptoms*	yes; *no*
├*offs duration*	yes; no
├*MentalProblems*	yes; *no*
├*age*	lt65; 65-75; gt75
IncreaseLDdosage	yes; *no*
├*usingLD*	yes; *no*
├*maxLD*	yes; no
├*MotorSymptoms*	yes; *no*
├*MentalProblems*	yes; *no*
├*dyskinesia duration*	yes; *no*
├*dyskinesia intensity*	yes; *no*
└*offs duration*	yes; no
DecreaseLDdosage	yes; *no*
├*usingLD*	yes; *no*
├*MotorSymptoms*	yes; *no*
├*MentalProblems*	yes; *no*
├*dyskinesia intensity*	yes; *no*
├*dyskinesia duration*	yes; *no*
└*offs duration*	yes; no
IncreaseLDintake	yes; *no*
├*usingLD*	yes; *no*
├*maxLD*	yes; no
├*MotorSymptoms*	yes; *no*
├*MentalProblems*	yes; *no*
├*dyskinesia intensity*	yes; *no*
├*dyskinesia duration*	yes; *no*
└*offs duration*	yes; no

DecreaseLDintake	yes; *no*
├*usingLD*	yes; *no*
├*MotorSymptoms*	yes; *no*
├*MentalProblems*	yes; *no*
├*dyskinesia intensity*	yes; *no*
├*dyskinesia duration*	yes; *no*
└*offs duration*	yes; no
ChangeDA+LDtoLD	yes; *no*
├*usingDA*	yes; *no*
├*usingLD*	yes; *no*
├*MentalProblems*	yes; *no*
└*cardiovascular*	yes; no
ChangeMAOItoMAOI+DA	yes; *no*
├*usingMAOI*	yes; *no*
├*MotorSymptoms*	yes; *no*
├*MentalProblems*	yes; *no*
ChangeMAOItoMAOI+LD	yes; *no*
├*usingMAOI*	yes; *no*
└*MotorSymptoms*	yes; *no*
StopMAOI	yes; *no*
├*usingMAOI*	yes; *no*
├*usingDA*	yes; *no*
├*Dyskinesia*	yes; *no*
├*MentalProblems*	yes; *no*
└*hypertension*	yes; no
AddMAOI	yes; *no*
├*usingMAOI*	yes; *no*
├*usingDA*	yes; *no*
├*usingLD*	yes; *no*
├*offs duration*	yes; no
└*MotorSymptoms*	yes; *no*

MotorSymptoms	yes; *no*
├rigidity	yes; *no*
├**Tremor**	yes; *no*
│├tremor at rest	yes; *no*
│├action tremor	yes; *no*
│└postural tremor	yes; *no*
└bradykinesia	yes; *no*
MentalProblems	yes; *no*
├impulsivity	yes; *no*
├cognition	yes; *no*
└**Psychosis**	yes; *no*
├hallucinations	yes; *no*
└paranoia	yes; *no*
Comorbidities	yes; *no*
├cardiovascular	yes; no
├low blood pressure	yes; no
└hypertension	yes; no
Dyskinesia	yes; *no*
├offs duration	yes; no
├dyskinesia intensity	yes; *no*
└dyskinesia duration	yes; *no*
PersonalCharacteristics	inactive; *active*
├age	lt65; 65-75; gt75
└activity	yes; no
CurrentTherapy	max; yes; *no*
├usingMAOI	yes; *no*
├usingDA	yes; *no*
├usingLD	yes; *no*
├maxDA	yes; no
└maxLD	yes; no

Examples of decision rules in DEX presented in a tabular form (2b)

	usingDA	Mental Problems	cardiovascular	low blood pressure	ChangeDAtoLD
	70%	10%	10%	10%	
1	yes	yes	*	*	yes
2	yes	*	yes	*	yes
3	yes	*	*	yes	yes
4	*	*no*	no	no	*no*
5	*no*	*	*	*	*no*

The change of medication treatment from DA to LD should happen only when the patient already takes DA. The change may take place in in three different cases: the patient has mental problems, cardiovascular problems, or low blood pressure. Otherwise, the change to LD should not happen.

	usingDA	maxDA	Motor Symptoms	offs duration	Mental Problems	cardiovascular	age	activity	Increase DAdosage
	18%	18%	6%	6%	18%	18%	9%	6%	
1	yes	no	yes	*	*no*	no	lt65	*	yes
2	yes	no	yes	*	*no*	no	<=65-75	yes	yes
3	yes	no	*	yes	*no*	no	lt65	*	yes
4	yes	no	*	yes	*no*	no	<=65-75	yes	yes
5	*	yes	*	*	*	*	*	*	*no*
6	*	*	*	*	yes	*	*	*	*no*
7	*	*	*	*	*	yes	*	*	*no*
8	*	*	*	*	*	*	>=65-75	no	*no*
9	*	*	*	*	*	*	gt75	*	*no*
10	*	*	*no*	no	*	*	*	*	*no*
11	*no*	*	*	*	*	*	*	*	*no*

Increase of DA dosage should happen only when the patient already takes DA and the maximal aloud dosage has not been reached. Dosage may be increased when the patient is younger than 75 years and has no mental or cardiovascular problems however experiences worsening of motor symptoms or worsening of offs duration. Otherwise, DA dosage should not be increased.

Examples of the used legend and symbols in the decision rules and state-transition submodels:

ChangeDAtoLD – change therapy from dopamine agonist to levodopa.

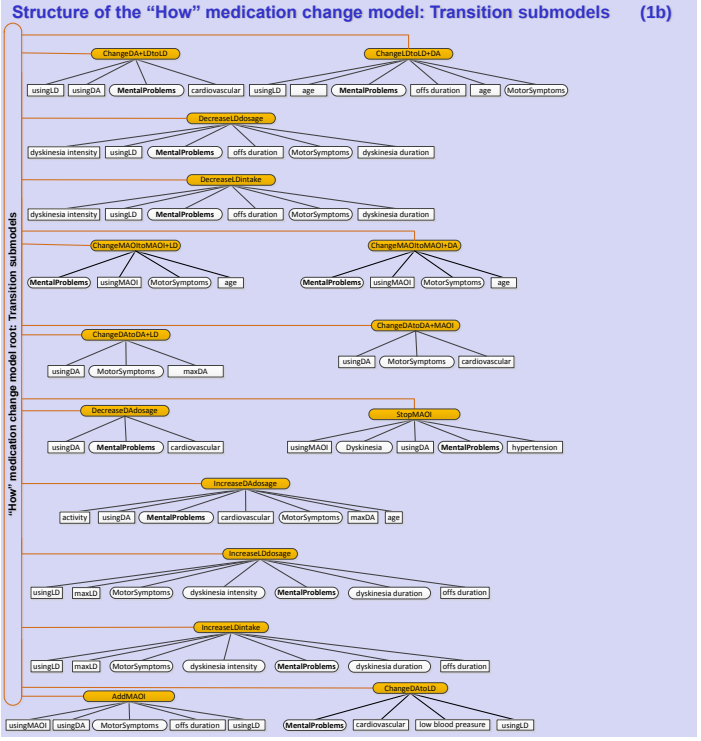
ChangeDAtoDA+MAOI – change therapy from dopamine agonist to dopamine agonist and MAO-B inhibitors.

IncreaseLDdosage – increase dosage of levodopa

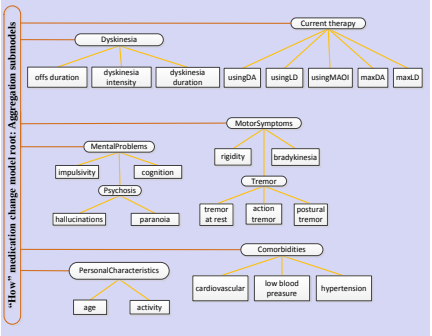
The symbol '*' used in the tables denotes any value that can appear at that position. For instance, in connection with an attribute than can take the values yes and no, the '*' stands for "yes or no".

The numbers below the heading denote approximate estimates of the strengths (weights) of the corresponding attribute; the larger the weight, the more important the attribute for the decision to change the medication.

Structure of the "How" medication change model: Transition submodels (1b)



Structure of the "How" medication change model: Aggregation submodels (1c)



Conclusions and future work

Using DEX methodology we developed a state-transition model for medication change of PD patients. This approach assured that the model is
- complete: it provides outputs for any possible inputs;
- robust: it works even if some input data is missing;
- consistent: the model is free of logical errors;
- transparent: the model is fully "open" for the inspection of contained decision rules;
- comprehensible: the embedded decision rules are easy to understand and explain.

The model has been evaluated on 2420 instances from PPMI dataset. Future work includes formal evaluation on real case patients scenarios and implementation in the PD_manager m-health platform.

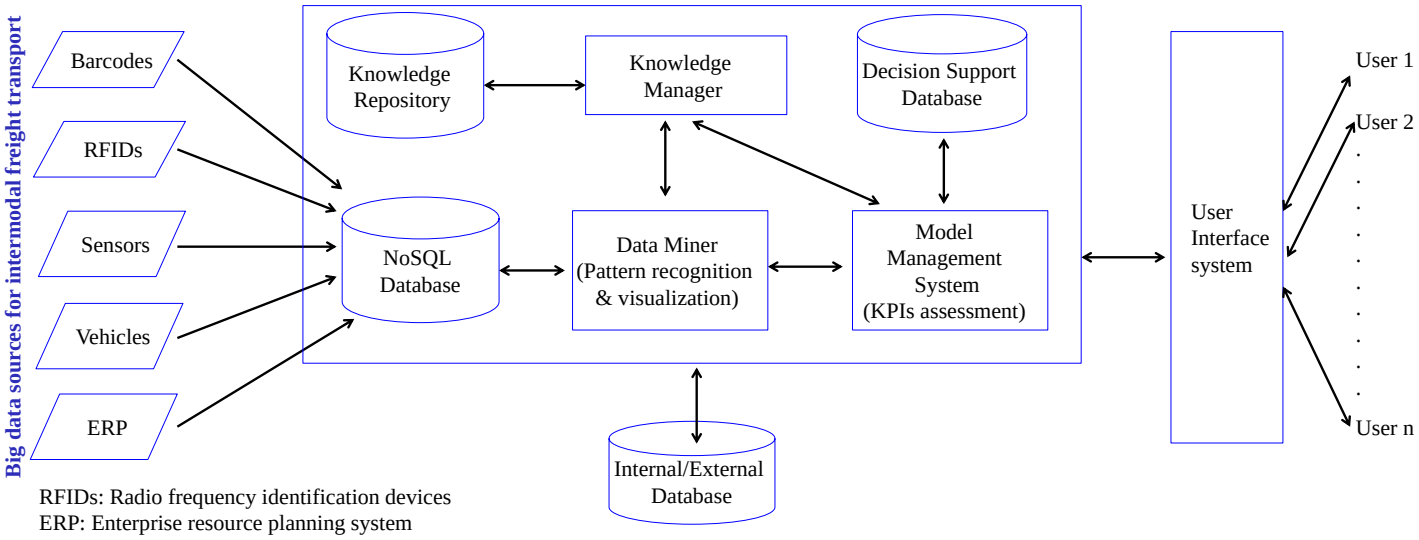
ICDSST 2017 Data, Information and Knowledge Visualisation in Decision Support Systems

An intelligent group decision support framework for sustainable performance management of intermodal freight transport

Fahham Hasan Qaiser, Alok Choudhary, Martin Sykora
School of Business and Economics
Loughborough University, Loughborough, UK
f.h.qaiser@lboro.ac.uk, a.choudhary@lboro.ac.uk, m.d.sykora@lboro.ac.uk

The proposed intelligent group decision support system (IGDSS) framework carries out sustainability assessment and management of intermodal freight transport system (IFTS). A set of key performance indicators (KPIs), covering economic, environmental and social dimensions of sustainability, is identified and selected through a separate research. The proposed system uses these KPIs to measure the performance of IFTS and identifies the potential areas where improvement can be made. In order to bring improvement in respective KPIs, a set of enablers are proposed by the system, which are certain factors, processes or actions that will help logistic companies to attain improvements in KPIs. These enablers will be stored in the knowledge repository of the system, which are identified and linked with respective KPIs through a separate research. While tapping the benefit from big data which the logistics companies are collecting with high volume, variety, velocity and veracity, this system attempts to discover interesting patterns, using data mining techniques, such as association and classification. The discovered patterns then transformed into knowledge, which can be visualized, thus supporting managers to make informed group decision making. This aids decision makers to understand the relationship among various data being collected as well as among KPIs, which helps predicting the value of a KPI in future while setting some input parameters. For example, the level of CO2 emission in future can be predicted by employees training, recycling and resource utilization, if data mining finds association rules among them. The discovered knowledge will then be shared, analysed and stored for future use.

A Simple Representation of Intelligent Group Decision Support System Architecture



RFIDs: Radio frequency identification devices
ERP: Enterprise resource planning system
NoSQL: A database used for wide variety of unstructured data

IGDSS ARCHITECTURE

The architecture is comprised of decision makers as users, user interface system, server, internal/external database and various sources of big data. Considering the characteristics of big data, a framework such as Hadoop may be used, which is capable of handling, processing and storing structured and unstructured data. The model management system (MMS) takes input from users and big data to measure the KPIs for IFTS. The MMS then co-ordinates with knowledge repository through knowledge manager to identify a set of enablers approprite to improve respective KPIs. Data miner also identify and visualize patterns among several parameters and KPIs to help managers making informed decision.

CONCLUSIONS & REMARKS

- This IGDSS enables a group of logistics managers to collaborate in a virtual environment and measure the performance of their intermodal freight transport system based on a set of KPIs.
- The system highlights the areas which needs attention for taking improvement actions. In this regard, the system proposes a set of enablers or drivers, which help bring KPIs up to the mark.
- Moreover, the system makes use of big data being collected from various sources to discover knowledge that would help managers to improve their system.

ICDSST 2017 Data, Information and Knowledge Visualisation in Decision Support Systems
Does the axial length of scatter plots influence decision making under risk?

Matthias Riekert, Prof. Dr. Sonja Utz, Prof. Dr. Katharina Scheiter
Leibniz-Institut für Wissensmedien
Schleichstraße 6, 72076 Tübingen, Germany
m.riekert@iwm-tuebingen.de, s.utz@iwm-tuebingen.de , k.scheiter@iwm-tuebingen.de

eurostars™
GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

Scatter plots are frequently used to visualize 2-dimensional data. They use a x- and y-axis to indicate the value of a specific data point. The length of the x- and y-axis is, however, variable and often depends on the screen size of the display device. For example, the y-axis is usually longer than the x-axis on mobile devices, but shorter on computer screens. In two experiments we investigated whether variations of axial length influence human decision-making (here: risk aversion). Participants had to choose between two lotteries. These lotteries were visualized as two data points on scatter plots that displayed the winning probability and the outcome on separate axes. We presented 30 trials, showing both one data point with low winning probability but high outcome (i.e. risky option) and one data point with high winning probability but low outcome (i.e. safe option). Based on previous literature, we hypothesized that extending the x-axis or y-axis increases the proportion of participants who choose the lottery with the higher value on the x-axis or y-axis respectively. Our results show that while participants were risk-averse and their choice was dependent on which lottery had the higher expected value, their choice was not influenced by the scatterplot's axial length. This result was independent of the axial assignment, which we changed between the two experiments (exp.1: N = 64; exp.2: N = 69).

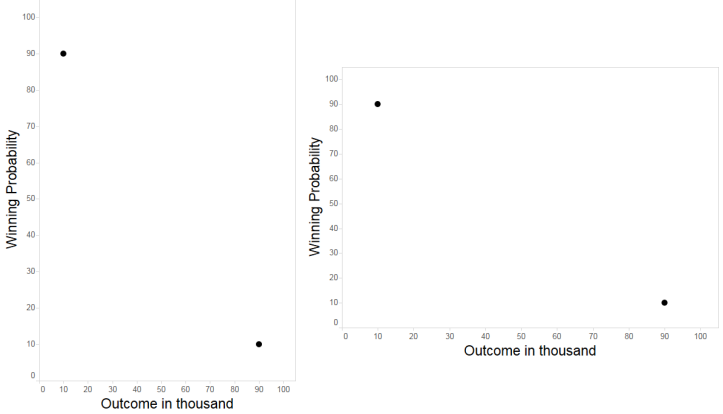
Introduction:

- Risk aversion is a human tendency in choices with positive outcomes [1]
- We investigated whether risk aversion can be reduced by manipulating the axial length of scatter plots.
- Based on previous research results, we expected that a data point with the highest value on one axis will become more salient by increasing the length of this axis [2]. This, in turn, increases its probability to be chosen [3].

Method:

- German speaking students were asked to make 30 decisions.
- We manipulated the axial length (longer x- or longer y- axis) and the expected value (EV) of the lotteries (Equal EV; safe option had higher EV; risky option had higher EV).
- Experiment 2 was identical, except that we reversed the axial assignment.

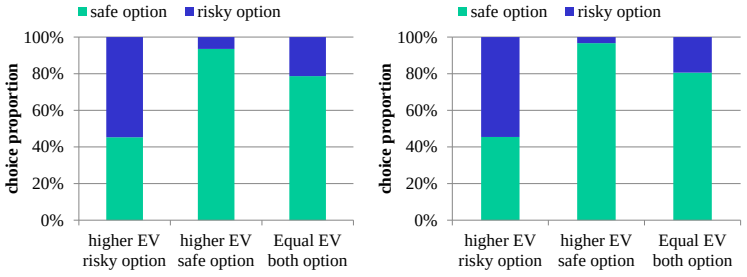
Illustration of the used material



Result:

- For both experiments, the statistical results were identical.
- Using a repeated-measure logistic regression, we found no interaction or main effect for axial length (all *p*s > .05), but a main effect for expected value.
- Post-hoc analyses showed that the choice proportion for the safe option was lowest when the riskier option's EV was higher, and highest when the safe option's EV was higher (all *p*s < .05).

Choice proportion for different expected values (exp. 1 and exp. 2)



CONCLUSIONS:

Our results suggests that varying the axial length is not a successful intervention when trying to reduce risk aversion. Nevertheless, other design features could influence people's risk aversion. Therefore, further research is needed to improve decision making using visualizations.

1. Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*.
2. Sun, Y., Li, S., & Bonini, N. (2010). Attribute salience in graphical representations affects evaluation. *Judgment and Decision Making*, 5(3), 151–158.
3. Dambacher, M., Haffke, P., Groß, D., & Hübner, R. (2015). Graphs versus numbers: How information format affects decisions in gambling. *Judgement and Decision Making*, *11*(3), 223–242.

ICDSST 2017 Data, Information and Knowledge Visualisation in Decision Support Systems
VISUAL ANALYTICS TOOLS FOR SOCIAL COGNOCRACY NETWORK

Alberto Turón, Alfredo Altuzarra, José María Moreno-Jiménez
Grupo Decisión Multicriterio Zaragoza. University of Zaragoza.
Gran Vía 2, 50005 Zaragoza, Spain
turon@unizar.es, altuzarr@unizar.es, moreno@unizar.es
web-page: http://gdmz.unizar.es

Social Cognocracy Network (SCN) is a collaborative social network developed by the Zaragoza Multicriteria Decision Making Group (GDMZ). SCN is based on the principles that support the cognitive democracy known as e-cognocracy. E-cognocracy uses two rounds in order to incorporate preferences through an e-voting module and an intermediate round of discussion in which the arguments that support the individual positions and decisions are added by means of a collaborative tool (forum). The forum provides quantitative measures of the actors' reputation and the relevance of topics and comments. Several visual analytics tools have been developed in order to analyze this quantitative information, so that social leaders -the persons whose opinions influence the preferences of others- can be identified. The visual tools have been codified as R functions, so that they can be integrated into interactive three-dimensional visualisation tools.

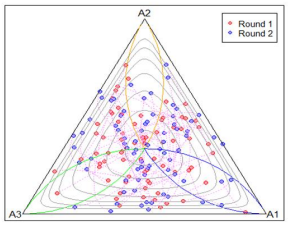
PREFERENCE STRUCTURES (2 rounds)

DM	A1	B1	C1	A2	B2	C2	Reputation
DM1	0,396	0,159	0,444	0,588	0,020	0,392	0,000
DM2	0,076	0,462	0,461	0,034	0,482	0,484	0,000
DM3	0,341	0,240	0,419	0,528	0,372	0,101	0,000
DM4	0,042	0,196	0,761	0,075	0,345	0,581	0,000
DM5	0,151	0,339	0,510	0,366	0,134	0,500	0,000
DM6	0,377	0,580	0,043	0,400	0,481	0,119	0,124
DM7	0,023	0,760	0,218	0,387	0,439	0,174	0,000
DM8	0,582	0,061	0,357	0,207	0,027	0,766	0,046
…	…	…	…	…	…	…	…

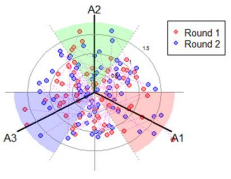
Preference structures are obtained by using AHP, integrated in the SCN

Three alternatives

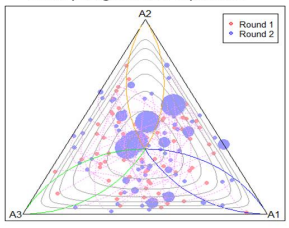
Ternary diagram



Euclidean ternary diagram



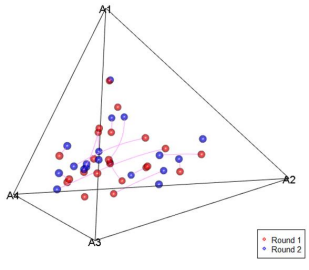
Ternary diagram with reputation



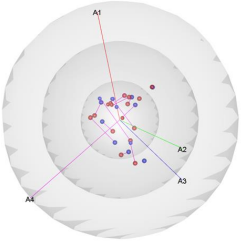
Ternary diagramas show the compositional distribution of the decision makers' priorities. The *Euclidean ternary diagram* maps the priorities in the euclidean space. The *Ternary diagram with reputation* shows the distribution of the most reputated decision makers (influencers).
Two-rounds ternary diagrams show the changes in the decision makers' preferences from one round to the next.

Four alternatives

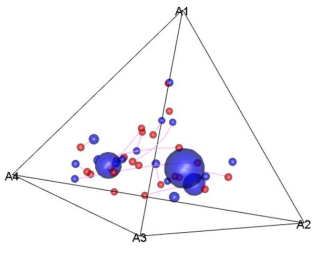
Quaternary diagram



Euclidean quaternary diagram



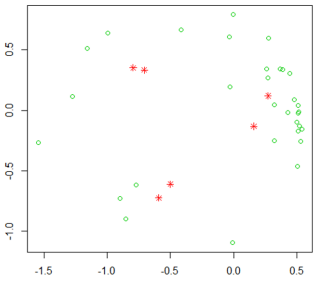
Quaternary diagram with reputation



Two-round quaternary diagrams. A decision problem with four alternatives can be explored in the simplex using an interactive 3D visualisation tool.

N alternatives

Multidimensional diagram



Problems with more than four alternatives can be represented in *Multidimensional diagrams* showing relative position of decision makers with respect to the alternatives.

- Preference structures are compositional data ($\sum_{i=1}^{n} w_i = 1$).
- Ternary diagrams project the preference structures in the compositional space.
- Three-dimensional interactive visualization is based in the *rgl* package (Adler *et al.*).
- In the Euclidean diagrams (isometric representation), distances represent euclidean distances between any pair of points.

SCN seeks the identification of social leaders, people whose opinions influence the determination of the structures of preference of citizens.

Visual analytics uses analytical reasoning through interactive visual interfaces in order to deal with problems whose size or complexity would be impossible to address by other procedures.

Our tools facilitate the visual analysis of the results of a decision making process and identify the influence of participants in the changes of preference structures.

This research was funded by the Spanish Ministry of Economy and Competitiveness along with FEDER funding (Project ECO2015-66673-R)

ICDSST 2017 Data, Information and Knowledge Visualisation in Decision Support Systems

Evaluation Assessment to Select Best Normalization Techniques for Multi-Criteria Decision Making Methods: ELECTRE, TOPSIS, SAW, VIKOR

Nazanin Vafaei, Rita A. Ribeiro, Luis M. Camarinha-Matos
UNINOVA-CTS and Faculty of Sciences and Technology, Nova University of Lisbon
2829-516 Caparica, Portugal
{Nazanin.vafaei, rar, cam}@uninova.pt
http://www.uninova.pt/

Introduction

Context:

- In most multi-criteria decision-making (MCDM) problems, criteria have different units.
- **Normalization** is a **transformation process** to obtain **numerical and comparable input data** by using a common scale.

Motivation:

- Importance of data normalization for decision problems
- Shortage of research studies available in this topic

RESEARCH QUESTION:

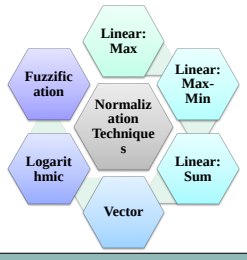
Which normalization technique is more suitable for usage with the MCDM methods: ELECTRE, TOPSIS, SAW, VIKOR?

Evaluation Assessment

Steps:

1) ***B&C-*** Calculate the difference between **Borda (number of wins) and copeland (number of losses) procedures** for alternatives in selected MCDM methods (Hwang and Yoon 1981).
2) ***Avg.-*** Calculating **Average rank of alternative** for the compared normalization techniques, per each MCDM method.
3) Comparative study between ranking of alternatives using **Pearson correlation** for Borda, copeland and Average values (Step1 & Step 2).
4) Calculating **Borda Count and Plurality voting (# wins)** for the three previous Steps (Srdjevic B. 2007).
5) Evaluating normalization techniques using *double preference* **ELECTRE** method.

More Common Normalization Techniques and MCDM Methods



Results of steps 1 to 3 for SAW method

Step 1 for SAW

	A1	A2	A3	A4	A5	A6	A7	B
A1		M	X	X	M	X	X	2
A2	X		X	X	M	X	X	1
A3	M	M		X	M	X	X	3
A4	M	M	X		M	M	M	5
A5	X	X	X	X		X	X	0
A6	X	X	X	X	M		X	1
A7	M	M	M	X	M	M		5
C	3	4	1	0	6	2	1	
B-C	-1	-3	2	5	-6	-1	4	
Rank	4	5	3	1	6	4	2	

Step 2 for SAW

	Max	Max-Min	Sum	Vector	Logarithmic	Fuzzification	Avg.	Rank
A1	1	4	1	1	2	5	2,333	1
A2	4	2	4	4	4	3	3,500	4
A3	3	1	3	3	5	1	2,667	2
A4	2	6	2	2	1	6	3,167	3
A5	5	3	5	5	3	2	3,833	5
A6	6	5	6	6	6	4	5,500	6
A7	7	7	7	7	7	7	7,000	7

Step 3 for SAW

	Max	Max-Min	Sum	Vector	Logarithmic	Fuzzification
Pearson with B & C	0,1347	-0,5837	0,1347	0,1347	0	-0,6286
Rank	1	3	1	1	2	4
Pearson with Avg.	0,9643	0,5000	0,9643	0,9643	0,6786	0,2857
Rank	1	5	1	1	4	6

Illustrative Example's Results for ELECTRE, TOPSIS, SAW, VIKOR

Step 4

		Max	Max-Min	Sum	Vector	Logarithmic	Fuzzification
Rank of NT * per each MCDM for B&C	ELECTRE	3	5	1	3	2	6
	SAW	1	5	1	1	4	6
	TOPSIS	3	5	2	4	1	6
	VIKOR	5	4	2	2	1	6
Borda count	sum	12	17	6	10	6	22
	B&C	12	7	18	14	18	2
	rank	4	5	1	3	1	6
Plurality voting	# Wins	1	0	2	1	2	0
	rank	3	5	1	3	1	5

* NT= Normalization Techniques

Step 5

	Max	Max-Min	Sum	Vector	Logarithmic	Fuzzification
Net dis-concordance	0,3548	2,8680	-3,8946	-1,7210	-2,6071	5
Rank	4	5	1	3	2	6

The priority of normalization technique's usage in the selected MCDM methods:

		Max	Max-Min	Linear-Sum	Vector	Logarithmic	Fuzzification
copeland & Borda Procedure	Borda count	4	5	1	3	1	6
	Plurality voting	3	5	1	3	1	5
ELECTRE	Double preference	4	5	1	2	3	6
Average	Borda count	1	4	3	1	5	6
	Plurality voting	2	3	3	1	5	5
Correlation	Borda count	1	3	4	2	6	5
	Plurality voting	1	2	4	2	4	4

Conclusion

- Based on the approaches from illustrative example **best** normalization technique for selected MCDM methods is **Linear sum,** and **Vector** and **Logarithmic** ones are in the next priority.
- The **worst** normalization techniques are **Fuzzification** and **Max-Min.**
- The results from **Average and Correlation metrics** are not discriminative enough, therefore copeland & Borda procedure are better as evaluation metrics.
- **ELECTRE** method **proved that the copeland & Borda procedure** ranking is trustable.