EmC-ICDSST 2019

5th International Conference on Decision Support System Technology – ICDSST 2019 & EURO Mini Conference 2019 on "Decision Support Systems: Main Developments & Future Trends"

University of Madeira, Portugal, 27th-29th May 2019

Editors: P. S. Abreu Freitas, F. Dargam, R., Ribeiro, J. M. Moreno Jimenez, J. Papathanasiou

ISBN: 978-989-8805-44-7

EWG DSS DECISION SUPPORT SYSTEMS

30 YEARS ANNIVERSARY

EDITORS

Paulo Sérgio Abreu Freitas
University of Madeira, Portugal

Fátima Dargam
SimTech Simulation Technology, Austria

Rita Ribeiro
UNINOVA University, Lisbon, Portugal

José Maria Moreno Jimenez
Zaragoza University, Spain

Jason Papathanasiou
University of Macedonia, Greece

Special Editorship Contribution

Pavlos Delias
International Hellenic University, Greece

Organization

Conference Organizing Chairs & Local Organizing Team

Paulo Sérgio Abreu Freitas
University of Madeira, Portugal
E-Mail: paulo.freitas@staff.uma.pt

Rita Ribeiro
UNINOVA University, Lisbon, Portugal
E-Mail: rar@uninova.pt

Fátima Dargam
SimTech Simulation Technology, Austria
E-Mail: f.dargam@simtechnology.com

Ana Luisa Respício
University of Lisbon, Portugal
E-Mail: respicio@di.fc.ul.pt

António Rodrigues
University of Lisbon, Portugal
E-Mail: ajrodrigues@fc.ul.pt

Jorge Freire de Souza
University of Porto, Portugal
E-Mail: jfsousa@fe.up.pt

Jorge Nélio Ferreira
University of Madeira, Portugal
E-Mail: jorge.nelio.ferreira@staff.uma.pt

Program Committee

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
Bertrand Mareschal	Université Libre de Bruxelles, Belgium
Boris Delibašić	University of Belgrade, Serbia
Carlos Henggeler Antunes	University of Coimbra, Portugal
Daouda Kamissoko	University of Toulouse, France
Dragana Bečejski-Vujaklija	Serbian Society for Informatics, Serbia
Fátima Dargam	SimTech Simulation Technology/ILTC, Austria
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
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
José Maria Moreno Jimenez	Zaragoza University, Spain
Kathrin Kirchner	Berlin School of Economics and Law, Germany
Lai Xu	Bournemouth University, UK
Marko Bohanec	Jozef Stefan Institute, Slovenia
Michail Madas	University of Macedonia, Greece
Milosz Kadzinski	Poznan University of Technology, Poland
Nikolaos Matsatsinis	Technical University of Crete, Greece
Nikolaos Ploskas	University of Macedonia, Greece
Panagiota Digkoglou	University of Macedonia, Greece
Pascale Zaraté	IRIT/Toulouse University, France
Pavlos Delias	International Hellenic University, Greece
Rita Ribeiro	UNINOVA – CA3, Portugal
Rudolf Vetschera	University of Vienna, Austria
Sandro Radovanović	University of Belgrade, Serbia
Shaofeng Liu	University of Plymouth, UK
Stefanos Tsiaras	Aristotle University of Thessaloniki, Greece
Theodore Tarnanidis	University of Macedonia, Greece
Uchitha Jayawickrama	Staffordshire University, UK
Wim Vanhoof	University of Namur, Belgium
Xin James He	Fairfield University, USA

Steering Committee – EWG-DSS Coordination Board

Pascale Zaraté	IRIT/Toulouse 1 Capitole University, France
Fátima Dargam	SimTech Simulation Technology, Austria

Shaofeng Liu	University of Plymouth, UK
Boris Delibašić	University of Belgrade, Serbia
Isabelle Linden	University of Namur, Belgium
Jason Papathanasiou	University of Macedonia, Greece
Pavlos Delias	International Hellenic University, Greece

Sponsors

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

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

APDIO - Associação Portuguesa de Investigação Operacional
(http://apdio.pt)

Institutional Sponsors

UMA Universidade da Madeira, Portugal
(http://www.uma.pt/)

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

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)

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

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

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

University of Macedonia, Department of Business Administration, Thessaloniki, Greece
(http://www.uom.gr/index.php?newlang=eng)

International Hellenic University, Greece
(https://www.ihu.edu.gr)

Industrial Sponsors

REACH Innovation EU
(http://www.reach-consultancy.eu/)

Deloitte
(https://www2.deloitte.com/)

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/)

Table of Contents

About the EWG-DSS

The EWG-DSS is a Working Group on Decision Support Systems within EURO, the Association of the European Operational Research Societies. The EWG-DSS was founded during a memorable EURO Summer Institute on DSS that took place at Madeira, Portugal, in May 1989. Most of the participants of the EURO Summer Institute on DSS in Madeira in 1989 still continue nowadays to pursue their goals, working actively in their research areas related to OR and Decision Support Systems.

The EWG-DSS was born with 24 founding-members. Since then, the number of members has substantially grown along the years. Now we are over 300 registered members coming from various nationalities. There has also been established quiet a few well-qualified research co-operations within the group members, which have generated valuable contributions to the DSS field in journal publications.

Since its creation, the EWG-DSS has held annual Meetings in various European countries, has taken active part in the EURO Conferences on decision-making related subjects; and has organized several workshops and conferences on different topics around Decision Support Systems.

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 professional academic and industrial opportunities for its members.
- Favour the development of innovative models, methods and tools in the field 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 streams in major conferences, as well as editing special and contributed issues in relevant scientific journals.

The process-loop shown next translates the main activities of the EWG-DSS envisaging the dissemination of DSS Information (1) and Research (2), in order to encourage DSS Development (3) and Collaboration (4) among the DSS researchers and professionals. As a consequence, Publication (5) opportunities to document the research & development processes and the end results are promoted within the EWG-DSS editions.

Specifically, to accomplish the main objectives listed above, the EWG-DSS promotes the following key-activities:

- Annual ICDSST Conference and other Conference-Streams organization related to Decision Support Systems topics.

- Annual Journal Special Issues publications, on support of the annual EWG-DSS organized Conferences, providing publication opportunities in the DSS Community.

- Annual EWG-DSS Newsletter publication, promoting the events and the research achievements of the EWG-DSS members and of the DSS Community as a whole.

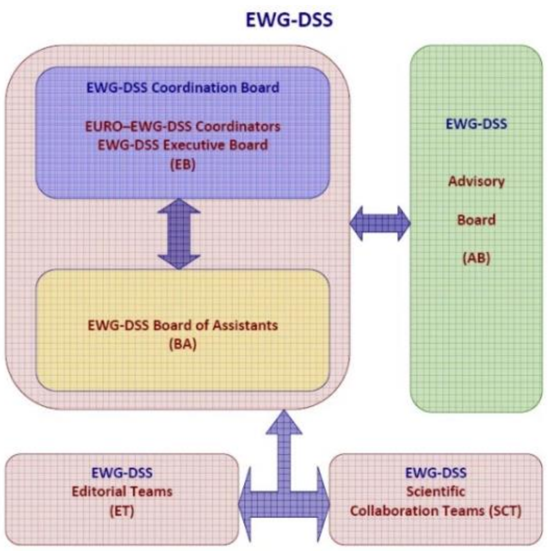
- Annual EWG-DSS-Award: a motivating research initiative for young researchers to submit and present their work in one of the EWG-DSS annual organized events.

- Collaboration projects among the group members. Check about the EWG-DSS research project Collab-Net Project, as well as the R&D competences of some member-institutions listed for collaboration in European projects.

Since 2007 the EWG-DSS has been managed by a Coordination Board. One of the aims of this coordination board is to promote joint-work among the group members and to encourage more participation of the whole group in DSS related projects and events, the best way possible. The Current EWG-DSS Coordination Board counts with the assistance of seven Board Members, namely: Prof. Pascale Zaraté, Dr. Fatima C.C. Dargam, Prof. Shaofeng Liu, Prof. Boris Delibasic, Prof. Isabelle Linden, Prof. Jason Papathanasiou and Prof. Pavlos Delias.

EWG-DSS Management Structure:

Since May 2015, the EWG-DSS has updated its Management Structure in order to incorporate to the Coordination Board (CB) an Advisory Board (AB), which is composed of senior members of the EWG-DSS group and of the international DSS community. The new group structure also distinguishes different teams, namely: the Scientific Collaboration Team (SCT) and the Editorial Team (ET), among the EWG-DSS Members who collaborate with EWG-DSS projects and joint–Editions, respectively. The figures below illustrate the current EWG-DSS Management Structure:



The Board of Assistants (BA) is formed by young researchers linked to the Coordination Board (CB) and Executive Board (EB) members, in order to assist them with the EWG-DSS annual tasks. Board and Team members can be in more than one of the groups of the EWG-DSS new defined Management Structure. However, the members of the Coordination Board are not supposed to be members of the Advisory Board and vice-versa.

The EWG-DSS Editorial Team (ET) is a dynamic group of researchers composed of EWG-DSS Members that have guest-edited Journal Special Issues and Springer Books with the EWG-DSS CB. The Scientific Collaboration Team (SCT) includes the researchers who are involved with EWG-DSS research projects and initiatives, like for instance the Collab-Net Project.

For more details about the EWG-DSS organized events and publications, check the homepage: http://ewgdss.wordpress.com/ .

Joining the EWG-DSS

The EWG-DSS membership does not cost you anything.
If you wish to join the EURO-Working Group on Decision Support Systems, all you have to do is to send an e-mail to our group at: <ewg-dss@fccdp.com>, with the following information:

Name; Affiliation; Mailing Address; Phone; e-mail; and Homepage link.

Alternatively, you can also join the EWG-DSS via our LinkedIn Group at:
http://www.linkedin.com/groups?about=&gid=1961459&trk=anet_ug_grppro

Thanks for your interest!
The EWG-DSS Coordination Board

Preface

ICDSST – the International Conference on Decision Support System Technology – is the flagship event of the Euro Working Group of Decision Support Systems (EWG-DSS). The ICDSST series of conference is relatively young and vibrant (since 2015), while its predecessor, including EWG-DSS workshops and summer schools, has a long tradition. The EWG-DSS was formally established in 1989 during a memorable EURO (European Association of Operational Research) summer school in Madeira, Portugal. Since then, the society has successfully organised wonderful events in many countries with collaboration from fantastic organisations, institutes and communities. This year's conference is a very special occasion – celebrating 30 years of fascinating journey, experience and achievements of the EWG-DSS, hence we feel privileged to have come back to the birthplace of the society with great support from EURO to hold this EmC-ICDSST 2019.

The EmC-ICDSST 2019 – EURO Mini (International) Conference on Decision Support System Technology has been organized by the EURO Working Group on Decision Support Systems EWG-DSS in cooperation with the Portuguese OR Association APDIO and the University of Madeira. The Conference is hold at the UMA University of Madeira, during the period of May 27th to May 29th, 2019.

The purpose of this EURO Mini Conference, focused on the main theme of "Decision Support Systems – Main Developments & Future Trends" is to bring together 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 DSS challenges in their respective areas of applications.

Researchers, engineers, computer scientists, OR and DSS professionals were encouraged to submit their work to the EmC-ICDSST 2019. The scientific topics of EmC-ICDSST 2019 include:

- advances in research on decision making and related areas
- artificial intelligence applied to decision support systems
- advances in applied decision support systems
- trends for new developments in decision support systems
- decision making integrated solutions within open data platforms
- knowledge management and resource discovery for decision making
- decision making methods, technologies and real-industry applications
- geographic information systems and decision making/support
- decision making, knowledge management and business intelligence
- DSS for business sustainability, innovation and entrepreneurship
- decision making in higher and school education

- innovative decision making approaches/methods and technologies
- big data analytics approaches for solving decision making issues
- big data visualization to support decision analysis and decision making
- social-networks analysis for decision making
- group and collaborative decision making
- multi-attribute and multi-criteria decision making
- approaches and advances in group decision making
- approaches and advances in negotiation decision support systems
- decision support systems and decision making in the health sector.

It is a great pleasure to introduce you to the papers and posters presented at EmC-ICDSST 2019, comprising this e-version of proceedings **"**Decision Support Systems: Main Developments & Future Trends"**.** Each paper/poster in the proceedings has been evaluated by multiple reviewers. The papers are published in two forms: a selection of high quality, long papers are published by Springer in their book series of "Lecture Notes in Business Information Processing", and short papers are published electronically. This year, we have introduced some new routines and tools to aid the search and citation of the papers, for example, all papers are assigned with a QR code. The papers have been classified into appropriate themes at conference presentation and in the proceedings. Hope you will find some really interesting information/knowledge and stimulating ideas within the broad area of decision making and decision support.

Apart from deploying the 5th EWG-DSS annual International Conference on Decision Support System Technology, the EmC-ICDSST 2019 event also has the objective of celebrating the 30th Anniversary of the EURO Working Group of Decision Support Systems EWG-DSS, created as a product of the ESI VI – EURO Summer Institute on DSS, in the Madeira Island in May 1989. Since then, the research area in DSS has substantially advanced and the EWG-DSS has consolidated its importance among the EURO Working Groups, by currently being one of its most stable and active groups considering its research initiatives and activities. The organization of the EmC-ICDSST 2019 in Madeira aims at recapitulating the developments of the Decision Support Systems area in the last 30 years, as well as enforcing the trends and new technologies in use, so that a consensus about the appropriate steps to be taken in future DSS research work can be established.

The conference takes place in the Rectory Building of the University of Madeira, UMA-Reitoria (Colégio dos Jesuítas) in the center of Funchal, Madeira, Portugal, during the period of May 27-29th, 2019. The beautiful Funchal town is situated on the wonderful Madeira Island, and is a tranquil and beautiful island in the middle of Atlantic Ocean, famous for its wine, sunshine, breath-taking sceneries, and wonderful food and restaurants. Nowadays Madeira is also the major touristic, cultural, artistic and politic center in the archipelago. We believe that in Madeira you can perfectly combine your work and leisure here at EmC-ICDSST 2019 by networking with some of the most respectable founding members of the EWG-DSS and world famous scholars/experts in the topic areas, and greatly enjoy the island at the same time! We look forward to facilitating insightful conversations which, without a doubt, will give rise to new and exciting ideas, approaches, technologies and projects for decision making and decision support!

Finally, we would like to thank many people who have greatly contributed to the success of this proceedings and the conference. Organizing a conference is certainly not an easy task and demands work, dedication and management efforts that we luckily find among the EWG-DSS Coordination Board members, Advisory Board members, Board of Assistants and members of the Organizing Committee of the EmC-ICDSST 2019. We sincerely wish to express our gratitude to all the helping hands that made this conference happen. The production of this volume would not have been possible without the precious support of Pavlos Delias and the excellent voluntary contributions of Sandro Radovanović, Panagiota Digkoglou, Maite Escobar Urmeneta, Jorge Freire de Sousa, Daouda Kamissoko, Pascale Zaraté, Boris Delibašić, and Shaofeng Liu, who worked in tandem with the Editors to perform a plethora of tasks, most of them emerging as urgent requirements. We are deeply thankful to them all! We also would like to thank the sponsorship and support from EURO, APDIO, and UMA for the realization of the conference. Last but not least, we wish to thank all authors and participants of the conference, for having trust and interest in the EWG-DSS organized conferences by submitting their work for reviewing, presentation and publication; and specially we need to warmly thank all Program Committee members for their excellent reviewing support on this conference. We have received 64 submissions and we had a tight schedule for supplying review feedbacks to the authors in time for the updates needed for publication and presentation. We would not have achieved our reviewing goals without the highly qualified, constructive and effective cooperation of our team of reviewers. Excellent job! ***THANKS to ALL of YOU!***

We wish all the EmC-ICDSST 2019 participants an enjoyable and fruitful collaboration time during the conference, with lots of networking for further cooperation and joint-work. Enjoy the EmC-ICDSST 2019! Enjoy Madeira!

The Editors:

Paulo Sérgio Abreu Freitas

Fátima Dargam

Rita Ribeiro

José Maria Moreno Jimenez

Jason Papathanasiou

Special Talks

Back to the Future in the Madeira Island?

José Paixão
Vice-Rector of the University of Lisbon, Portugal
Honorary Chair of the EmC-ICDSST 2019
Local organizer of the EURO Summer Institute on DSS in Madeira in 1989

Thirty years ago, a group of young researchers from 16 different nationalities gathered at Madeira Island for a Summer Institute on Decision Support Systems (DSS), promoted by EURO (The Association of European Operational Research Societies). The subject of the meeting was extremely challenging for the operational research community since, at the time, practitioners had difficulties to distinct a decision support system (DSS) from an expert systems (ES), an artificial intelligence technique. The purpose of the Institute was to bring together a group of young scientists for contributing to the clarification of the concept of DSS from an Operational Research perspective. Eventually, making clearer the distinction between DSS and ES.

Today, I can say that much more than that has been achieved as it is well proven by the vitality of the EWG-DSS, a joint venture of the "Madeira Group" that became one of the most active working groups in the EURO. In a simple tribute, I will try to outline the aspects that were fundamental for the success of the meeting held 30 years ago in Madeira. I believe that some of them still remain current and may be useful to achieve the objectives of ICDSST 2019.In fact, research on DSS cannot ignore the vertiginous technology wave that we are seeing today and that poses questions that go beyond the technical and conceptual aspects dominant 30 years ago. Inspired by the famous Zemeckis' movie, I am strongly convinced that new "young McFly's" will come to Madeira Island, enjoy the nice local atmosphere and pave the way for future developments on DSS research.

Decision Support Systems - Historical Innovations and Modern Technology Challenges

Christer Carlsson
Director of the Institute of Advanced Management Systems Research
Professor of Management Science of the Åbo Akademi University
Turku, Finland

Managerial tasks carry latent needs for support to do a better job; classical DSS had at its core the approach to support, not replace. We worked out a DSS called Woodstrat for strategic planning and management and could verify – in full-scale implementation - the DSS characteristics Sprague worked out and most of the DSS benefits Keen had found. This brings out the historical innovations that DSS introduced. We also found that a relevant and useful DSS could help "self-confident professionals" to back away from predictions on future demand and competition that could not find support in facts. The digital disruption of the 2010'es brought big data and the need for decision making in almost real time. It also introduced analytics and faster, more effective algorithms developed in computational intelligence. The ICT of the 2010'es invites decision-making anywhere and anytime which is a challenge for the cognitive ability of decision makers. The road map for DSS for the 2020'es points to digital coaching systems that adapt to the cognitive levels of the users.

For a full version of this talk, please see this page, at the end of this book.

DSS Approaches in Mobility & Societal Issues

A decision support system for urban transport planning based on dynamics, correlations and causalities among road paths

Glykeria Myrovali, Panagiotis Tzenos, Maria Morfoulaki, Georgia Aifadopoulou
Hellenic Institute of Transport/Centre for Research and Technology Hellas (HIT/CERTH),
6th km Charilaou-Thermi Road, Thessaloniki, 57001, Greece
{myrovali, ptzenos, marmor,gea}@certh.gr

Theodoros Karakasidis, Avraam Charakopoulos
Laboratory of Hydromechanics and Environmental Engineering, Department of Civil Engineering, University of Thessaly, 38334 Volos, Greece
thkarak@uth.gr, avracha@yahoo.gr

ABSTRACT

The great abundance of multi-sensor traffic data (traditional traffic data sources - loops, cameras and radars accompanied or even replaced by the most recent - Bluetooth detectors, GPS enabled floating car data) although offering the chance to exploit Big Data advantages in traffic planning, management and monitoring, has also opened the debate on data cleaning, fusion and interpretation techniques. The current paper concentrates on floating taxi data in the case of a Greek city, Thessaloniki city, and proposes the use of advanced spatiotemporal dynamics identification techniques among urban road paths for gaining a deep understanding of complex relations among them. The visualizations deriving from the advanced time series analysis proposed (hereinafter referred also as knowledge graphs) facilitate the understanding of the relations and the potential future reactions/outcomes of urban traffic management and calming interventions, enhances communication potentials (useful and consumable by any target group) and therefore add on the acceptability and effectiveness of decision making. The paper concludes in the proposal of an abstract Decision Support System to forecast, predict or potentially preempt any negative outcomes that could come from not looking directly to long datasets.

Keywords: Big Traffic Data, Cross correlation, Granger causality, Decision Support Systems, Mobility Patterns, travel time, floating taxi data

Find and cite the full work as:
Myrovali,G., Karakasidis, T., Charakopoulos, A., Tzenos, P., Morfoulaki M., Aifadopoulou , G., 2019. A decision support system for urban transport planning based on dynamics, correlations and causalities among road paths, in: Freitas, P.S.A., Dargam, F., Moreno, J.M. (Eds.), Decision Support Systems IX: DSS Developments & Future Trends, Lecture Notes in Business Information Processing. Springer International Publishing, Cham, pp. 26–38.

Facility Location Problem in Public Safety Area

Ana Paula Henriques de Gusmão, Rafaella Maria Aragão Pereira, Maisa Mendonça Silva, Bruno Ferreira da Costa Borba
Research Group in Information and Decision Systems (GPSID) – Universidade Federal de Pernambuco (UFPE), Av. da Arquitetura, Cidade Universitária, 50740-550, Recife - PE, Brazil
anapaulagusmao@cdsid.com; maisa@cdsid.org.br; rafaellamaragao@gmail.com; brunoborba50@hotmail.com

ABSTRACT

This paper aims to support decisions related to determining efficient spatial distributions of police units, given that the location of these units is a strategic matter regarding the costs of operations and police response times to occurrences. On conducting a literature review of the facility location problem in the public security area, it was noticed the mathematical model Maximal Covering Location Problem (MCLP) could be applied to a case study concerning military police units in Recife, a large city in Northeast Brazil. For the case study, MCLP, along with a Decision Support System (DSS), was applied to make an analysis of the potential location of military police facilities in Recife. This evaluation enabled the results obtained from an ideal scenario, where the bases were positioned at optimal points, to be compared with the performance of these facilities in their current location. For this purpose, CVP (an indicator for Violent Crimes against property georeferenced data from 2017 and the location of the military police units of Recife were used.

Keywords: Public Security, Facility Location Problem, Decision Support System.

Find and cite the full work as:
Henriques de Gusmão, A.P., Aragão Pereira, R.M., Silva, M.M., Costa Borba, B.F., 2019. Facility Location Problem in Public Safety Area, in: Freitas, P.S.A., Dargam, F., Moreno, J.M. (Eds.), Decision Support Systems IX: DSS Developments & Future Trends, Lecture Notes in Business Information Processing. Springer International Publishing, Cham, pp. 13-25.

Forecasting Emotions in Refugee and Migrant Communities for Site Management Support Using Fuzzy Cognitive Maps

Maria Drakaki, Panagiotis Tzionas
Department of Automation Engineering, Alexander Technological Educational Institute of Thessaloniki, P.O. Box 141, GR-574 00, Thessaloniki, Hellas (GR)
{drakaki, ptzionas}@autom.teithe.gr

Hacer Güner Gören
Department of Industrial Engineering, Pamukkale University,
Kinikli Campus, Denizli, Turkey
hgoren@pau.edu.tr

ABSTRACT

Refugees and migrants arrivals in the Mediterranean since 2014 have resulted in humanitarian relief operations at national and international levels. Decision making by relevant actors in all aspects of humanitarian response impacts on the well-being of People of Concern (PoC). Site management decisions take into account a wide range of criteria in diverse sectors, including protection of PoC. Criteria address basic relief response, such as water, sanitation and nutrition, social characteristics, as well as protection of PoC. In particular, site management decisions may affect relationships between PoC, leading to tensions or peaceful coexistence between PoC. In this paper a decision making process is proposed for the forecasting of emotions in PoC sites in order to assist site management decisions, particularly in the context of avoidance of tensions. The method uses Fuzzy Cognitive Maps (FCM) to forecast emotions based on input data obtained from site management reports. Historical data from site management reports in Greece have been used for deriving the indicators used in the input layer of the designed FCM. The proposed method will be applied to forecast potential tensions in PoC sites in Greece.

Keywords: Artificial emotion forecasting; fuzzy cognitive maps; refugee site management; decision support method; refugee crisis.

Find and cite the full work as:
Drakaki, M., Gören, H.G., Tzionas, P., 2019. Forecasting Emotions in Refugee and Migrant Communities for Site Management Support Using Fuzzy Cognitive Maps, in: Freitas, P.S.A., Dargam, F., Moreno, J.M. (Eds.), Decision Support Systems IX: DSS Developments & Future Trends, Lecture Notes in Business Information Processing. Springer International Publishing, Cham, pp. 1–12.

A Conceptual Framework for an Integrated Information System to enhance Urban Mobility

Sérgio-Pedro Duarte, Jorge Freire de Sousa, Jorge Pinho de Sousa
INESC TEC / Faculty of Engineering, University of Porto
Rua Dr. Roberto Frias, 4200-265 Porto, Portugal
s.duarte@fe.up.pt, jfsousa@fe.up.pt, jorge.p.sousa@inesctec.pt

ABSTRACT

Decision-making processes are essential in the design and operation of transportation systems. Multiple actors interact through these systems and their decisions may significantly influence each other actions, both at a strategic level and an operational and real time level. However, those decisions are often based on the same or similar data. Planners are interested in knowing congestion points to manage traffic, transport operators want to prevent delays, and travelers try to choose the best route for their trips.

In this context, multimodal networks have a great potential to provide higher levels of service, and to improve system efficiency and city sustainability. Managing a multimodal network with multiple stakeholders can be achieved through integration and collaboration, and a smart incorporation of modes, along with well-designed interfaces and multi-platform information systems. In such a context, a diverse and dynamic demand from varied types of travelers (residents, visitors) and trip purposes (loading/unloading, services access) needs to be addressed. In cargo transport, vehicle dimensions, pollutant emissions or time/zone windows have also to be taken into account.

To help decision-making processes in such multi-stakeholder contexts, we have created a conceptual framework for an integrated information system based on Enterprise Architecture Frameworks and Service Design approaches. Since this information system will be capable of integrating different data sources and provide different information for different stakeholders, the framework can hopefully be used in quite different contexts. In this work, we use it to address urban mobility issues.

Keywords: urban mobility, smart sustainable cities, integrated information systems, enterprise architecture frameworks, multi-user platforms, decision-making.

INTRODUCTION

Urban population growth is causing new challenges to transportation systems efficiency. In the EU alone, urban population may reach 80% of the total population by 2020. Statistics show that urban areas are growing, leading to more mega cities with larger suburbs [1]. Due to socio-demographic changes related to technological developments, we are facing a socio-technical transition that will significantly change citizens behavior regarding transportation systems [2]. Whether it is a vehicle related technology, a set of gadgets or a simple smartphone app, current technology is changing citizens' habits and activities. This transition environment puts greater pressure in the different decision-making processes. In the perspective of authorities or operators, they need to be confident that their decisions will benefit the overall system efficiency; and in the perspective of the users, they need to plan their trips in the most efficient way according to their personal needs and possibly conflicting criteria. Therefore, the complexity of information and its impacts on the different stakeholders require a more advanced and integrated approach.

Such integration can be achieved by an information system (IS) capable of responding to different types of decisions and taking into account different stakeholders' needs [3]. Nevertheless, the development of an integrated system raises challenges due to the existence of data from multiple sources, that needs to be properly managed and analyzed [4, 5]. One integrated approach will increase the participation of all stakeholders, in the sense that users can feed the system with data that other stakeholders can use. In that way, such a system would also improve planning activities and bring citizens to participate in designing a smart, sustainable city through knowledge co-creation [6].

Designing and implementing an integrated IS should be based on a framework that considers the different user perspectives, the different levels of decision and the different type of data. As far as we know, existing approaches are based on Enterprise Architecture Frameworks (EAF), such as the Zachman Framework for Enterprise Architecture (or the Zachman Framework) [7]. However, recently proposed guidelines can be used to build integrated information systems in several other contexts [8]. Zachman himself recognizes that no architecture is completely right or completely wrong. Therefore they need to be used together [9]. In this paper, we present a conceptual framework for the design of an IS for urban mobility issues.

We argue that to deliver an IS that supports the decision processes of the different stakeholders, there is a need to properly relate all types of possible decisions to the required data. Hence, in our approach we integrate methods from Service Design in the Zachman Framework, thus adding value to the current state of the art on the mobility research. That helps us to study the complexity of the service ecosystem in the context of urban transport. The next section presents a brief literature review on EAF and Multilevel Service Design (MSD). The conceptual framework developed in this work is presented in another section of the paper. Finally, some conclusions and future research topics are presented.

ENTERPRISE ARCHITECTURE FRAMEWORKS AND SERVICE DESIGN

Enterprise Architecture (EA) first appeared as a form of aligning IS with companies' processes by using information technology (IT) as a facilitator of those processes. EA can be seen as a set of guidelines, methods, models, and tools that can be used to design an IS structure that supports an organizational structure [10].

From the existing EAF, we base our research on the Zachman Framework and the Enterprise Architecture Design (EAD) from Velho [11]. EAD is based on The Open Group Architecture Framework (TOGAF), Federal Enterprise Architecture Framework (FEAF), and Enterprise Architecture Planning (EAP) [11]. Regarding Service Design approaches, we use the Multilevel Service Design proposed by Patrício et al. [12].

The Zachman Framework for Enterprise Architecture can be considered as the first EA. This framework is based on comparing the construction of a IS to the construction of a building. The original framework presented five perspectives, from the owner of the project to the sub-contractor, including planner, designer, and builder, like in physical construction. For each of these perspectives, the framework proposes the documentation to be produced at different abstractions (data, function, network, people, time, and motivation) [7, 11].

This type of organization gives the Zachman Framework the potential for being well understood by cross-functional teams. It is represented as a matrix in which the perspective of each player can be associated to one row, and each fundamental question to one column. Each player perspective is associated with a level of detail, with this level growing when moving in the matrix, from top to bottom [9]. The level of detail is related to the type of document produced, with a clear distinction between business processes, entities, application, and technology. The Zachman Framework was recently modified to include the user perspective, represented by the role of workers in construction, thus including the description of the operations to be performed [13]. This helps using IS as a service since it includes both backend and frontend perspectives.

EAD emphasizes the planning phase of the project since it defines the strategy for the IS to be implemented. Thus, it can also be studied as a methodology. It presents the information system design in four major phases (planning; current situation; situation to be implemented; and implementation plan and change management) [11].

Moreover, EAD highlights the importance of project management by including policy layers (Management Principles, and Governance Models), and a transition plan [11]. This transition plan is included because EAD considers the existence of a previous architecture and system. This is more accurate in the XXI century since most companies already have some IS supporting their processes. In the same way as the Zachman Framework, EAD uses the business model process (BPM) methods to map the processes to be supported by the information system.

Multilevel Service Design (MSD) is a method to design technology-enabled services focused on the customer perspective. It was created to respond to the challenges of multi-interface services, e.g., services where customers need to use physical and digital channels [14]. MSD is an interdisciplinary method or process that approaches services design from a high level (the service concept) to more detailed specifications (the service encounter), including the definition of customers interactions moments (the service system). MSD takes into account customer’s requirements, enabling the integration of the customer experience in all the stages. Moreover, it does not only consider the customer perspective, but it also enables the customer’s participation in the design process, by using the appropriated methods [12]. Some of the MSD features are related to Business Process Management (BPM), but they have been adapted to increase awareness about customer experience.

CONCEPTUAL FRAMEWORK

The conceptual framework proposed in this work is built on the idea that the IS should provide the right information to support decision-making, and that the collected data should provide such information. By first mapping the decisions for the different stakeholders, we can then map the information provided by the data, to support the various decision processes.

The IS architecture is designed in three layers (Acquisition, Processing, Visualization) that correspond to the first three stages of the decision-making process (Data collection, Data Analysis, and Results Analysis).

As in some other EAF, the first level of the framework is concerned with the definition of the goals. In our work, this takes place during the process of mapping the decisions, as this process will help define the real purpose of the IS. Then, identifying all the data sources will allow us to build the data acquisition layer. For this purpose, the adopted methodology is based on approaching the problem in the opposite way of the decision-making process. This means that identifying the data sources required can be achieved by answering the following questions:

1. What decision will the user make?
2. What information is required to make that decision?
3. What data can provide that information?

Then, the process goes on with the design of the other layers of the system (Figure 1). At the visualization layer, different applications may be developed to respond to each stakeholder decision-making process.

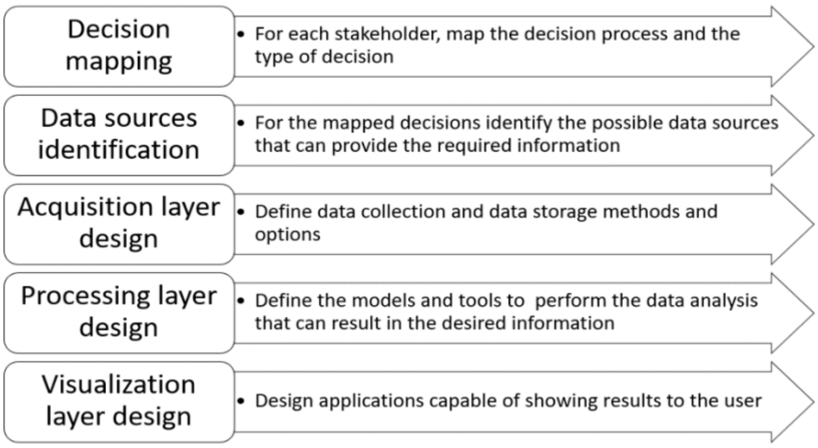


Figure 1. Conceptual framework.

The foundations of our work are similar to those of other IS frameworks, as it includes an objective (defining what the system is supposed to do), a model (how the system will work and interact with the users) and a technological architecture (where the system will run). Due to the complexity of the transportation sector and the need to involve multiple stakeholders, we integrate Multilevel Service Design [12] concepts in the process. Figure 2 demonstrates how our framework is related to the Zachman Framework.

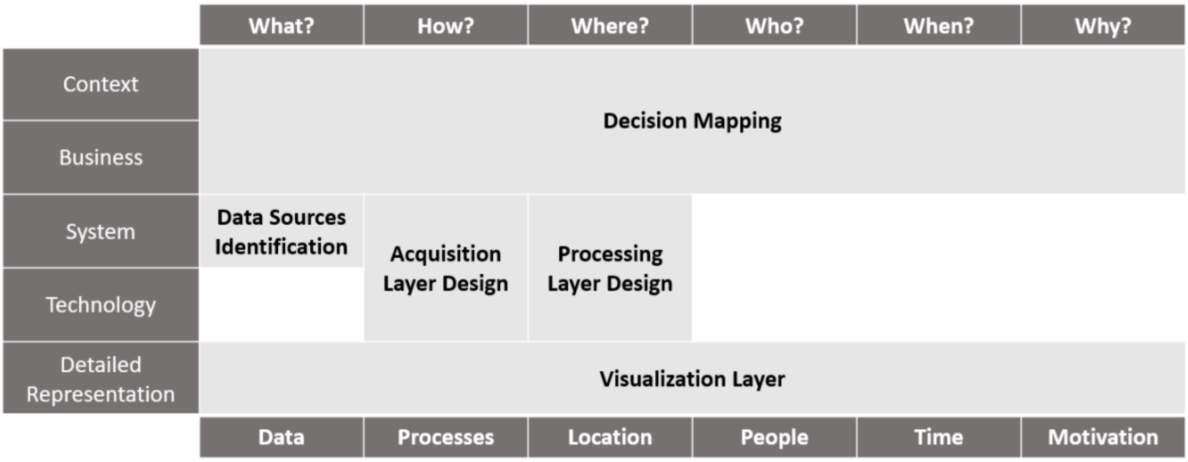


Figure 2. Relation of our conceptual framework to the Zachman Framework.

The decision mapping phase is related to the Service Concept from MSD and the Strategic-Operational Summary from EAD. By mapping the decisions of the stakeholders, we also analyze the stakeholders' profiles. This helps in defining the mission and objectives of the IS as EAD suggests. At the Context level, studying the Value Constellation helps us understand the ecosystem and the relations between the stakeholders. This is important because, in the transportation sector, the decisions of each stakeholder can influence other stakeholders. For instance, the citizens' behavior influences the operators' supply, as the existing supply influences the citizens' choice of using or not using public transport. At the Business level, we use artifacts from MSD such as System Architecture and System Navigation to answer the *what?*, *how?* and *who?* questions from the Zachman's abstractions. At this point, we map the decision process for each stakeholder (*who*) by defining *what* their actions are and *how* they currently structure their decisions. Regarding the design of the Acquisition and Processing layers, it is related to the data, application and technology architectures from EAD. Finally, the visualization layer details, for each stakeholder, the interaction with the IS. This can be achieved using the service experience blueprint from MSD.

CONCLUSIONS

Our research is built on two main premises: i) a multilevel integrated planning process for multimodal mobility is key to increase the sustainability of a city; ii) well-informed decisions may have significant positive impacts. Accordingly, an integrated IS should allow all stakeholders to have access to the same information when making their decisions, whether they are complex and strategic (e.g. adding a new road) or simple (such as planning a trip). The participation of multiple stakeholders is one of the main challenges, and it is the main reason why we use Service Design concepts to develop a user focused approach.

The main contributions of this work cover both the development of a IS architecture and the stakeholder involvement. The support provided to the stakeholder inclusive decision processes is based on existing methodologies, but it is enhanced by putting together concepts of Service Design with EAF. This will also raise the awareness of the service perspective in city planning. Once there are multiple stakeholders and several ways to interact with transport systems, the complexity of such systems resembles the complexity of a service with multiple interaction channels. In our case, the IS resulting from the proposed architecture allows service providers to improve their service and communication with the transportation system users.

As future research, new practical stakeholder involvement methods can be developed, mainly using ICT to improve communication, collaboration, and transparency among the different stakeholders. This will be achieved during the decision-mapping phase, where a participatory process will help to validate decisions.

ACKNOWLEDGEMENTS

This work is partially funded by Fundação para a Ciência e Tecnologia (FCT), Portugal, through grant PD/BD/128064/2016.

REFERENCES

[1] ALICE, "Urban Freight: research and innovation roadmap," p. 60, 2015.
[2] A. Spickermann, V. Grienitz, and H. A. Von Der Gracht, "Heading towards a multimodal city of the future: Multi-stakeholder scenarios for urban mobility," *Technol. Forecast. Soc. Change*, vol. 89, pp. 201–221, 2014.
[3] Z. Zhang, "Healthcare Information System Architecture Design Based on Big Data," vol. 118, no. Amcce, pp. 61–64, 2017.
[4] P. Rajapaksha, R. Farahbakhsh, E. Nathanail, and N. Crespi, "iTrip , a framework to enhance urban mobility by leveraging various data sources," *Transp. Res. Procedia*, vol. 24, no. 2016, pp. 113–122, 2017.
[5] M. C. Beutel, S. Gökay, W. Kluth, K. Krempels, C. Samsel, and C. Terwelp, "Product Oriented Integration of Heterogeneous Mobility Services," pp. 1529–1534, 2014.
[6] A. Deligiannidou and D. Amaxilatis, "Knowledge Co-creation in the OrganiCity : Data Annotation with JAMAiCA," pp. 717–722, 2016.
[7] S. Mrdalj, "A comparison of enterprise architecture frameworks," no. January 2006, 2016.
[8] W. Hasselbring, "Information System Integration," *Commun. ACM*, vol. 43 no 6, no. June, pp. 32–38, 2000.
[9] R. Sessions, "A comparison of the top four enterprise-architecture methodologies," *Houst. ObjectWatch Inc*, 2007.
[10] J. Danny, Shanlunt, G. Wang, and H. Alianto, "The Application of Zachman Framework in Improving Better Decision Making," *1st 2018 Indones. Assoc. Pattern Recognit. Int. Conf. Ina. 2018 - Proc.*, pp. 245–249, 2019.
[11] A. V. Velho, *Arquitectura de empresas*. Centro Atlântico, 2004.
[12] L. Patrício, R. P. Fisk, J. F. e Cunha, and L. Constantine, "Multilevel service design: From customer value constellation to service experience blueprinting," *J. Serv. Res.*, vol. 14, no. 2, pp. 180–200, 2011.
[13] J. P. Zachman, "The Zachman framework evolution" *Zachman Int. Monum. CO*, 2009.
[14] L. Patrício, R. P. Fisk, and J. Falcão e Cunha, "Designing multi-interface service experiences: The service experience blueprint," *J. Serv. Res.*, vol. 10, no. 4, pp. 318–334, 2008.

DSS Methodologies & Techniques

A Decision Support System for resilience based on functionality analysis of interconnected systems

Daouda Kamissoko, Frederick Benaben, Amendeep Amendeep
Centre Génie Industriel, IMT Mines Albi, University of Toulouse
Campus Jarlard, 81013 Albi CT Cédex 09
daouda.kamissoko@mines-albi.fr, frederick.benaben@mines-albi.fr, aamandee@mines-albi.fr

Blazho Nastov
Axellience, Lille
4 Avenue des Saules, 59160 Lille
blazho.nastov@axellience.fr

Vincent Chapurlat, Aurelia Bony-Dandrieuxd, Jérôme Tixier, Nicolas Daclin
IMT Mines Alès, University of Montpellier
6 Avenue de Clavières, 30100 Alès
vincent.chapurlat@mines-ales.fr, aurelia.bony-dandrieux@mines-ales.fr,
jerome.tixier@mines-ales.fr, nicolas.daclin@mines-ales.fr

ABSTRACT

The increasing number of disruptions to Critical Infrastructure, like natural disasters, terrorist attacks, or internal failure is today a major problem of society. Concern is even greater when considering the interconnected nature of Critical Infrastructure, which might lead to failure propagation, causing domino and cascade effects. To mitigate such outcomes, critical infrastructure must recover its capacity to function with regard to several criteria. Stakeholders must therefore analyse and improve the resilience of critical infrastructure before any disruption occurs, and base this analysis on different models so as to guarantee society's vital needs. Current resilience assessment methods are mainly oriented towards the context of a single system, thus narrowing their criteria metrics, limiting flexibility and adaptation to other contexts, and overlooking the interconnected nature of systems. This article introduces a new Decision Support System that makes it possible to define a model to evaluate the functionalities of interconnected systems. The model is then used to assess the resilience of these systems based on simple and generic criteria that can be extended and adapted.

Keywords: Resilience; Functionality evaluation; Resilience assessment; Risk, Decision; Criteria; Simulation, Critical Infrastructure, System

INTRODUCTION

Modern society relies on the functioning and mutual exchange of services of various interconnected and interdependent infrastructures, i.e., systems or systems of systems (healthcare, energy, transport, manufacturing, financial, etc.) [1], [2]. The interdependencies between and within systems make them less resilient to disruptions [3]. Specifically, a disruption in an interconnected system can lead to domino and cascade effect that impacts on the other systems related to the initially affected one [4]. This becomes a problem when considering the currently increasing number of natural disasters. For instance, in July 2012, the largest blackout in history affected more than 600 million people in India. Through a cascade effect, several other systems (transport, telecommunication, finance…) also failed [5], [6]. In 2011, the flooding in South-East Asia led to a lack of hard drives and to an increase in the price of these devices all over the world [7]. The growing number of hurricanes - Sandy, Isabel, Harvey and Irma - has provoked not only human and material damage, but also economic and production/service capacity failures [8].

The notion of resilience is related to the functioning of critical infrastructures or systems and is here defined as "the capacity of a system to recover, in a minimum time, with minimum costs (financial, human, workload, etc.) a certain functioning capacity on all dimensions of its performances". Some aspects of the resilience of a system can be assessed by analysing its functionalities in several situations: (a) before a disruptive event, (b) during a disruptive event and (c) after a disruptive event. During each situation, it is important to (i) be able to assess the resilience at a given timestamp and/or period, including during disruptive events, and (ii) to identify preventive actions for different scenarios, to improve the results of the resilience assessment.

This paper focuses on situation (a) and proposes a Decision Support System for continuous and multidimensional resilience assessment based on analysis of the functionalities of interconnected systems.

Current resilience assessment approaches are oriented towards individual systems [9], whether they be financial, healthcare or transport systems. These approaches are therefore inflexible (difficult to adapt to other domains), with fixed criteria that generally concern performance (other criteria that might be important in the assessment of resilience are overlooked) and not applicable in the context of interconnected systems. This paper defines Decision Support System that makes it possible to define a model to evaluate the functionalities of interconnected systems that are linked to some aspects of resilience. The model is further used to assess resilience, based on several criteria. The proposed generic criteria can easily be extended and adapted depending on the context and needs. The originality of the approach lies in (1) the combination of functionality-analysis models and continuous resilience assessment following several dimensions of systems, (2) flexibility of criteria metrics for easy adaptation in different contexts and (3) the possibility to aggregate the results of several functionality-analysis models with continuous assessment of the resilience of interconnected systems.

The paper is organized as follows. The contribution is presented in the second section and illustrated in the third section by a case study, before concluding in the last section by an evaluation of the limits and perspectives of the research work.

Proposal

The main concepts of the proposal are illustrated in Figure 3. A system can be composed of several components. A territory can hosts several systems. The relations between systems might be functional or non-functional. Functional relations allow the circulation of flows. Non-functional relations refer to influence relations described by [10]. A component is characterized by several criteria. A criterion is defined by a value and a unit. To consider uncertainty, in normal functioning, a criterion value is comprised between a minimum value f_{min}^{n} and a maximum one f_{max}^{n}. The value of a criterion must not be outside the limits f_{max} and f_{min}. Criteria and Systems are composable (i.e., contain other criteria or systems). An evolution function defines a behaviour to simulate the evolution of a criterion and to change its value over time. The aggregation function, as its name suggests, aggregates the values of at least two criteria. The influence function determines how a given criterion/system changes another one under specified conditions. Note that feared events and flows are also instances of the concept System (see Figure 1).

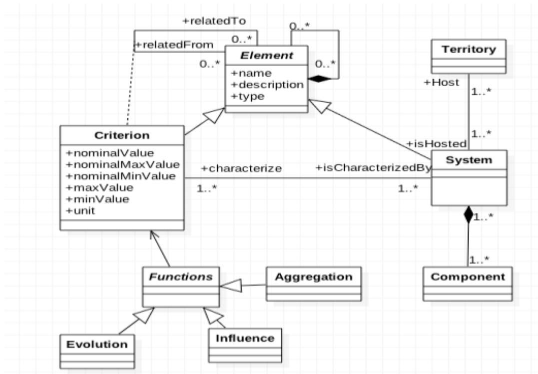


Figure 1. Conceptual model.

A functionality evaluation process defines a Functionality Evaluation Model that is further used for resilience assessment. This process is based on three steps: 1) modelling, 2) model transformation and 3) simulation. Step 1 – modelling consists of identifying systems, criteria, evolution, aggregation and influence functions. After that, relevant criteria characterizing the system are also identified. Then a function that manages every criterion is defined if needed. From a technical point of view, the modelling of these concepts is done in the Obeo Designer environment (https://www.obeodesigner.com/en). Step 2 – model transformation is an automated process that transforms designed models into a simulation environment. Technically, a model-to-text transformation script is written in Obeo Aceleo to perform the transformation from Obeo Designer to the GAMA platform. This transformation is based on [11] and [12]. Step 3 – simulation consists of simulating the results of the model transformation and plotting the result (i.e., a functionality curve). This is technically provided through the GAMA multi-agent simulation platform (http://gama-platform.org/).

This proposal is included in the framework of the MAIIEUTIC project (http://maiieutic.mines-ales.fr/) founded by the CARNOT M.I.N.E.S institute. The validation process includes several meetings with experts from System Engineering, Crisis Management, Risk Management and Multi-criteria Decision-Making approaches.

Throughout this paper, assertions 1 to 10 below define some aspects of the Functionality Evaluation Model related to resilience assessment.

- Assertion (1): Several elementary independent criteria that characterize a system can be used to define a Functionality Evaluation Model as a parameter to assess a particular aspect of the resilience of this system.

- Assertion (2): A Functionality Evaluation Model must consider the objectives and constraints of the Territory.

- Assertion (3): The value of the assessed resilience is between 0 (not resilient at all) and 1 (fully resilient).

- Assertion (4): The value of assessed resilience of a system depends on the value of each criterion aggregated into the Functionality Evaluation Model.

Assertion (5): Based on the Functionality Evaluation Model, the value of a given criterion must not be greater than certain values.

- Assertion (6): Based on the Functionality Evaluation Model, the value of a given criterion must be between certain values.

Assertion (7): There is a decrease in the value of assessed resilience due to the value of the criteria dropping out of the limits f_{min}^n and f_{max}^n.

Assertion (8): Based on a Functionality Evaluation Model, the closer the value of criteria is to the limit f_{max} or f_{min} the lower the value of assessed resilience is.

Assertion (9): Based on a Functionality Evaluation Model, if the value of a criterion is between f_{min} and f_{min}^n or f_{max} and f_{max}^n (this situation is denoted as "the system stays in a bad functioning period") for more than a given period, then the value of assessed resilience decreases.

Assertion (10): The more often the value of a criterion goes outside of f_{min}^n and f_{max}^n, the lower the value of the assessed resilience is.

Case study

The proposed methodology is applied to a simple case study. The aim here is to demonstrate the claims of the proposal. The system considered here is a network of infrastructures composed of: (1) a coal power plant, (2) a road (transport system), and (3) a signalling system. The power plant produces a certain quantity of electricity and needs trucks to deliver coal. The amount of electricity depends on the overall production capacity and the quantity of coal in incoming trucks (i.e., the quantity of transported coal). The performance of the signalling system depends on the electricity used. A bad performance impacts the safety of the road that furthermore influences the number of trucks in circulation (i.e., the quantity of transported coal to produce electricity). Therefore, a decreasing number of trucks decreases the electricity production of the power plant. Figure 2 shows the relationship between the criteria of these systems, where the arrows represent the dependence relationship.

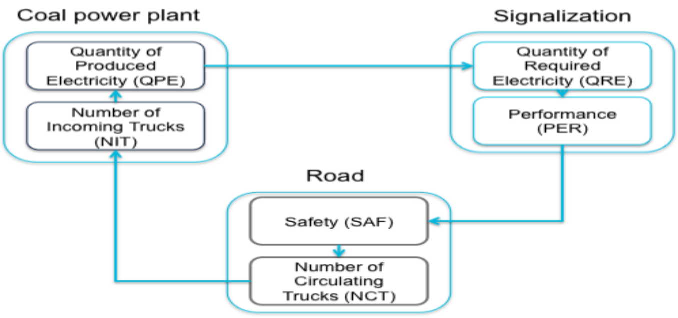


Figure 2. A use case based on a network of infrastructures.

This example is modelled using the Obeo Designer platform, transformed based on an Acceleo transformation script and simulated based on the multi-agent platform GAMA. The focus here is on the simulation results.

There are six criteria: The Quantity of Produced Electricity (QPE), the Number of Incoming Trucks (NIT), the Quantity of Required Electricity (QRE), the Performance (PER), the Safety (SAF) and the Number of Circulating Trucks (NCT). These criteria characterize the global system, as shown in the Figure 7. The proposal complies with Assertions 1 to 4; 6, defined in the previous section. For the sake of space and simplicity, only the Quantity of Produced Electricity (QPE) criterion is discussed hereafter. The objective of the power plant, assumed for this case study, is the production of 5.0 units (e.g., gigawatts); authorized minimum and maximum are assumed to be respectively 0.1 and 10. Note that Figure 3 defines the metamodel for these parameters. In normal functioning, the fluctuation might be between [3.0 - 7.0]. 5 units are tolerated if the objective is not fulfilled. The evolution function for the QPE, i.e., $f_n(t)$ is defined below.

$$\begin{gathered} f_n(0) = 5.0 \\ f_n(t+1) = f_n(t) \pm \rho(0.5) \end{gathered} \tag{1}$$

Where $\rho(x)$ is a random function that for a given number x returns a random number n such that $0 \leq n \leq x$.

In addition, the value of f_{min}^n, f_{max}^n, f_{min} and f_{max} should change with respect to time. For instance, the power consumption of a city depends on seasons (e.g., electricity consumption increases during winter) and thus the minimal production of the power plant should be greater than 3.0. Consequently, the evolution functions defined by Equation (25 – 28) are assigned to f_{min}^n, f_{max}^n, f_{min} and f_{max}. Note that for the sake of simplicity, these functions randomly change the value of their parameters.

$$\begin{gathered} f_{min}(0) = 0.1 \\ f_{min}(t+1) = f_{min}(t) \pm \rho(0.1) \end{gathered} \tag{2}$$

$$\begin{gathered} f_{max}(0) = 10 \\ f_{max}(t+1) = f_{max}(t) \pm \rho(0.1) \end{gathered} \tag{3}$$

$$\begin{gathered} f_{min}^n(0) = 3.0 \\ f_{min}^n(t+1) = f_{min}^n(t) \pm \rho(0.1) \end{gathered} \tag{4}$$

$$f^n_{max}(0) = 7.0$$
$$f^n_{max}(t+1) = f^n_{max}(t) \pm \rho(0.1) \quad (5)$$

For all situations the value of assessed resilience is between 0 and 1 and depends on all six criteria. However, there is no constraint on the duration of the bad functioning periods, and consequently, Assertions 9 and 10 are not tested in this use case. For the other assertions, i.e., Assertions 5; 7 and 8, two simulations are proposed: one for the up value, and one for the down value. For each assertion, the simulation result shown here is for 12 timestamps. For example, the Figure 8 illustrates a disruptive event that causes a drop-up (overload) of the QPE that cannot be mitigated and causes a system crash. Resilience factor R^{up}_{C0} remains 1 while the value of the criteria is superior to f_{min} or inferior to f_{max}, until the disruptive event occurs at timestamp 9 (Figure 8). After timestamp 9, a drop-up of the QPE causes the system to enter a bad functioning period. Shortly after, i.e., just before timestamp 10, the system enters an unacceptable state because the QPE becomes superior to f_{max} and R^{up}_{C0} becomes 0

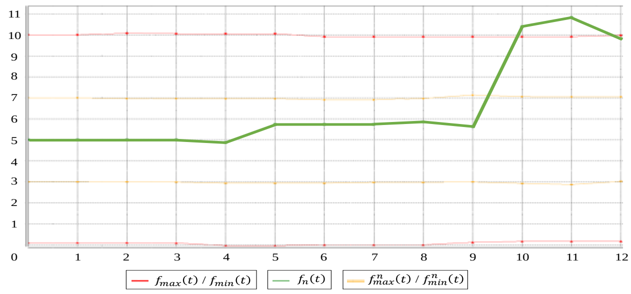


Figure 8: Simulation

Conlusion

Infrastructures are nowadays facing an increasing number of disruptions, from natural disasters to terrorist attacks and internal failures. Mitigating these negative effects means that the infrastructures must recover their initial functioning capacity in minimum time, with minimum costs, i.e., they must be resilient. Thus, stakeholders must analyse the resilience of infrastructures before any disruption to anticipate the right decisions at the right time. This paper introduces a Decision Support System for the evaluation of the functionality of interconnected systems for the purpose of resilience assessment. The methodology is designed to prepare Stakeholders for different disruptions before they happen, allowing them to make the right decisions at the right time. The major contributions of the methodology are:

- It provides a relevant tool for the combination of functionality analysis models and continuous resilience assessment following several dimensions of systems;
- It proposes a way to connect rough data with functional analysis and resilience assessment;
- It proposes an agile, evolutive and continuous resilience assessment paradigm (it can be extended with other criteria, other data sources to define criteria, other aggregation formulas, etc.).

The main limitation of the approach is its dependence on relevant input data. In other words, an incorrect domain knowledge as input will surely provide an incorrect evaluation. In this sense, we are unable to warn users about the inconsistency of their input data.

REFERENCES

[1] M. W. Maier, *The Art of Systems Architecting, Third Edition*. CRC Press, 2009.

[2] L.-G. Mattsson et E. Jenelius, « Vulnerability and resilience of transport systems – A discussion of recent research », *Transp. Res. Part Policy Pract.*, vol. 81, p. 16-34, nov. 2015.

[3] J. Johanson, « Risk and vulnérability Analysis of Interdependent Technical Infrastructures », Lund University, Dept. of Measurement Technology and Industrial Electrical Engineering, 2010.

[4] M. Ouyang, L. Dueñas-Osorio, et X. Min, « A three-stage resilience analysis framework for urban infrastructure systems », *Struct. Saf.*, vol. 36-37, p. 23-31, mai 2012.

[5] L. L. Lai, H. T. Zhang, C. S. Lai, F. Y. Xu, et S. Mishra, « Investigation on July 2012 Indian blackout », in *2013 International Conference on Machine Learning and Cybernetics*, 2013, vol. 01, p. 92-97.

[6] J. J. Romero, « Blackouts illuminate India's power problems », *IEEE Spectr.*, vol. 49, nº 10, p. 11-12, oct. 2012.

[7] C. Booth, *Strategic Procurement: Organizing Suppliers and Supply Chains for Competitive Advantage*. Kogan Page Publishers, 2014.

[8] F. Saleh, V. Ramaswamy, Y. Wang, N. Georgas, A. Blumberg, et J. Pullen, « A multi-scale ensemble-based framework for forecasting compound coastal-riverine flooding: The Hackensack-Passaic watershed and Newark Bay », *Adv. Water Resour.*, vol. 110, p. 371-386, déc. 2017.

[9] C. Pursiainen, « Critical infrastructure resilience: A Nordic model in the making? », *Int. J. Disaster Risk Reduct.*, vol. 27, p. 632-641, mars 2018.

[10] D. Kamissoko, « Decision support for infrastructure network vulnerability assessment in natural disaster crisis situations », Doctorate, University of Toulouse, University of Toulouse 1 Capitole, 2013.

[11] B. Nastov, V. Chapurlat, C. Dony, et F. Pfister, « A Tooled Approach for Designing Executable and Verifiable Modeling Languages », *INSIGHT Q. Mag. Int. Counc. Syst. Eng. INCOSE*, vol. 18, nº 4, p. 31–33, 2016.

[12] B. Nastov, V. Chapurlat, Dony, Christophe, et F. Pfister, « Towards V&V suitable Domain Specific Modeling Languages for MBSE: A tooled approach », in *the 26th Annual INCOSE International Symposium (IS'2016)*, 2016.

Maturity models in project management for software development: a literature review

João Batista Sarmento dos Santos Neto[1,2], Ana Paula Cabral Seixas Costa[2]

[1]Universidade Federal de Mato Grosso do Sul
Cidade Universitária, S/N, Campo Grande, Mato Grosso do Sul State, Brazil
joao.sarmento@ufms.br
web-page: https://faeng.ufms.br/graduacao/bacharelado/engenharia-de-producao/

[2]Universidade Federal de Pernambuco
Av. Prof. Moraes Rego, 1235, Cidade Universitária, Recife, 50670-901, Pernambuco State, Brazil,
apcabral@cdsid.org.br
web-page: https://www.ufpe.br/dep

ABSTRACT

In a scenario where organizations in the software development industry face a highly competitive environment, there is a need to find ways to support decision-making to satisfy customers and achieve business excellence. A maturity model is regarded as an important tool for demonstrating the evolution of a given approach and allows organizations to engage on planning actions to achieve the desired results. Thus, this study aims to investigate which maturity models are being applied in the domain of project management for software development. To achieve this goal, a literature review was carried out based on a review of 36 articles. The research produced evidence that showed that although many maturity models for project management are discussed, few such models are applied in the software development segment. It is believed that this occurs due to two maturity models, CMM and CMMI, which have been widely discussed in the literature, having become firmly established.

Keywords: Maturity Model, Project Management, Software Engineering, Capability Maturity Model Integration, CMMI.

INTRODUCTION

Maturity models (MMs) are considered management tools that help a decision-maker to define the best way to improve a specific domain i.e. MMs tend to be used in managing organizational change. When deciding which domain to improve, a number of different directions may be taken. Often, it is not easy to decide how to modify the process in the best possible way [1]. This is one of the reasons why MMs are used during the improvement process.

Although MMs first appeared in the 70s to control Information Systems, they have been used in different domains [2]. One of the areas that gained prominence was project management. Managing projects is a difficult activity for any organization. Various events can lead to failure, such as non-compliance with deadlines, a poorly planned budget, or nonconformity of predefined specifications [3]. It can be said that the main challenge of project management is to achieve all the goals and objectives of the project while adhering to the main constraints placed on it, such as scope, quality, deadlines and budget.

In 1987, Carnegie Mellon University's Software Engineering Institute (SEI) took the lead by proposing the Capability Maturity Model (CMM), an MM from the point of view of the processes of software engineering. The model offers five levels of maturity to help software companies continually improve their software processing capabilities [4]. The successful application of CMM in the software industry has inspired experts internationally from many fields, like the well-discussed research on the development of a project management maturity model. As a result, many valuable MMs in project management have been used ever since [5]. Thus, the main objective of this paper was to identify which MMs are being applied in the software development segment, for use in the field of project management. For this, a literature review was carried out, which took advantage of the final set of relevant articles identified by the systematic review of literature conducted by [2]. The intention is that the present study could be used as an aid for software development organizations whose desire to assess the level of their project maturity, however, not knowing which MM to choose. This work belongs to a wider study that seeks for MM evaluation for selection problem.

METHODOLOGY

According to [6], a research study always starts by setting out to tackle some problem, and does so by asking a question. Therefore, the problem that drives this study is: Which MMs are being applied to the project management domain in the software development industry? To answer this question, an exploratory-descriptive study was conducted based on a literature review. The papers to be reviewed were selected from the study by [2]. For their research, the objective of the systematic review of the literature undertaken by these authors was to examine published papers that presented the development of the research area in MMs, as well as evaluation models for classifying maturity levels. Figure 1 illustrates the research model and the steps followed to carry out this study.

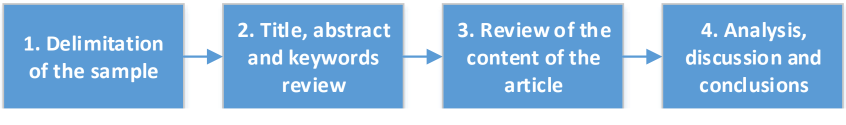


Figure 3: Research model

In the final set of relevant articles selected by [2], 409 articles were identified for more than 20 different domains. For this paper, the sample selected from those in [2] was limited to articles identified as being for “Project management” and “Software engineering/IT” or “IS management”. In addition, the authors of [2] classified the selected articles according to their research focus: application, development, meta-analysis and validation. The objective of this study is to define only the MMs that have been applied in some organization. Thus, a sample of 36 articles was selected (Appendix A). In step 2 of the research model, the titles, abstracts and keywords of all articles delimited in the sampling were reviewed. The aim of this step was to select only MMs that had been applied in project management that guided the segment of software development. Articles focusing on other segments were discarded from the analysis. In step 3, review of the content of the article, each article was read. If in step 2 it was not possible to identify if the revised article should be part of the analysis, in step 3, the article was read thoroughly, thereby ensuring that only papers relevant to this study were selected. Step 4 was used to analyze and discuss the findings, and, presents the conclusions. For the analysis, a classification scheme was defined with the purpose of facilitating the comparison and discussion of the articles selected. The classification scheme is shown in Table 1.

Table 1: Classification scheme

Classes	Examples
MMs applied	CMM, CMMI, OPM3, PMMM, K-PMMM
Procedure for setting the maturity level	Sets a score, Fuzzy Number
Approach of the Paper	Descriptive, Prescriptive or Benchmarking

RESULTS AND DISCUSSION

By applying the research model, 36 articles for the selected domains were identified. Of these, 15 articles were selected for the domain of "Project management" and 21 for "Software engineering/IT or IS management". Using the VOSviewer® tool enabled a bibliometric map of the selected articles in the sample to be generated (Figure 2). The bibliometric map constructed displayed the density of occurrence of keywords in the titles and abstracts. On the maps that were generated of keywords or clusters of greatest recurrence, a greater or lesser distance between items represents the relative strength between these words [7]. The greater the number of items near a point and the greater the proximity of neighboring items, the more strength the keyword has, or the higher the density [2]. In the bibliometric map of Figure 2 each color represents a different cluster.

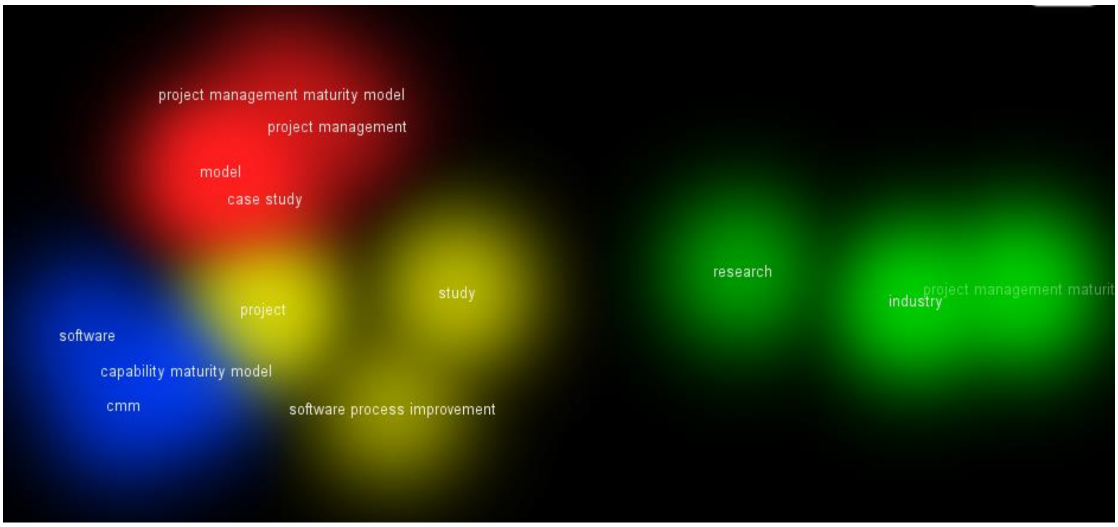


Figure 2: A view of the density in the clusters of the most frequent keywords

It was verified that the keywords ‘model’, ‘project’, ‘project management maturity’ and ‘capability maturity model’ are the words that had the highest density, i.e. the greatest relevance of occurrence. This suggests that research on these topics is more advanced. Note the proximity of some of these terms in the formation of 4 clusters with items concentrated around these keywords (Figure 2). What is noticeable in Figure 2 is the proximity of the red, blue and yellow clusters, that is, the clusters that have as the main keyword the words ‘model’, ‘capability maturity model’ and ‘project’ respectively. This may be an indication that, for these areas of study, besides being mature areas of research, their performance was close to each other in the studies carried out. In other words, these terms are interrelated in academic studies.

By proceeding with steps 2 and 3 of the research model, two articles that contribute in a relevant way to this research [8, 9] were identified. Both articles selected discussed the application of an MM in the software segment with a project management approach. They were selected from the “Software engineering/IT or IS management” domain. In other words, the articles of the domain "Project management" have application in other industrial segments. Of the two articles selected, both discuss the application of CMMI (Capability Maturity Model Integration). The CMMI are models created by the SEI and are already well established in the MM literature. The CMM is based on the principles and practices of total quality originally developed by [10]. From the CMM, the SEI launched the CMMI in 2001 to integrate the then already existing CMMs [11]. The CMMI program describes how to evolve an immature process into a mature, disciplined process while taking into account the domains of process management, project management, engineering, and support [11]. Some other MMs were identified in the research; however, without their having been applied in software development. Table 2 shows the relationship of MMs in the management of the projects which were identified from the sample of selected articles. Also shown in Table 2 are the maturity levels of each model, whether elements of the PMBOK have been incorporated in its structure, to what segment the application belongs according to the article and the main advantages on the MMs.

As to the evaluation models used to evaluate the level of maturity in the two publications, it was observed in the studies developed in [8, 9], Fuzzy Theory was used to deal with linguistic terms and to define the level of maturity after applying the CMMI. Fuzzy theory has techniques that are designed to deal with imprecise linguistic concepts or fuzzy terms. It allows users to enter inaccurate terms and yet get more consistent answers [2].

Table 2: Maturity models in the project management domain

MM	Maturity level	Incorporation of PMBOK elements	Example of the application	Main advantages
Kerzner’s Project Maturity Model (KPMMM)	Level 1- Common Language, Level 2- Common Processes, Level 3- Singular Methodology, Level 4- Benchmarking and Level 5- Continuous Improvement	Yes	Consultants; contractors; manufacturers; service providers; investment companies and financial institutions	It has an assessment software with free online version and an application questionnaire in the guide: Using the Project Management Maturity Model. The model offers some references on how to build an improvement plan based on results.
PM Solutions’ Project	Level 1- Initial Process, Level 2- Structured Process and Standards, Level 3- Organizational	Yes	professional, scientific and technical services;	It has an assessment software for PM Solutions consultants. Has an application questionnaire. Has a book/ guide from

Management Maturity Model (PMMM)	and Institutionalized Standards, Level 4- Managed Process and Level 5- Optimizing Process		information; finance and insurance; manufacturing; mining and engineering industry; construction industries; public agencies	orientation. The assessed results are used as a base for the improvement of the plan development, however, the model does not show how to do it.
Organizational Project Management Maturity Model (OPM3)	Level 1- Standardize, Level 2- Measure, Level 3- Control and Level 4- Improve	Yes	Construction industries; various (without specifying which segments)	It has an evaluation tool and the improvement plan is one of the key elements of the model. Includes the project, program and portfolio dimensions in its structure. Has a book/ guide for orientation.
PRINCE2 Capability Maturity Model (P2CMM)	Level 1- Cognitive, Level 2- Repeatable, Level 3- Management, Level 4- Repeatable and Level 5- Cognitive	No	Simulation	It has a free practices guide. Includes the project, program and portfolio dimensions in its structure. It is based on the PRINCE2 and P3M3 project management method. The assessed results are used as a base for the improvement of the plan development, however, the model does not show how to do it.
Prado Project Management Maturity Model (P2M3)	Level 1- Initial (ad hoc), Level 2- Known, Level 3- Standardized, Level 4- Managed and Level 5- Optimized	Yes	R&D companies	It has a free evaluation tool. Includes the project, program and portfolio dimensions in its structure. The assessed results are used as a base for the improvement of the plan development, however, the model does not show how to do it. Has a book/ guide for orientation.
Portfolio, Program and Project Management Maturity Model (P3M3)	Level 1- Awareness of process, Level 2- Repeatable process, Level 3- Defined process, Level 4- managed process and Level 5- Optimized process	Yes	Government agencies	Includes the project, program and portfolio dimensions in its structure. Has a guide for orientation. The assessed results are used as a base for the improvement of the plan development, however, the model does not show how to do it.
Capability Maturity Model	Level 1- Initial, Level 2- Managed, Level 3- Defined, Level 4-	No	Software engineering	It has an evaluation method for professionals authorized by SEI. Has a guide for orientation. Has two ways to demonstrating

Integration (CMMI)	Quantitatively Managed and Level 5- Optimizing			maturity: continuous and staged representation. The assessed results are used as a base for the improvement of the plan development.

Besides classifying the procedures used to define the level of maturity, the authors of this article sought to determine whether the applications took a descriptive, prescriptive or benchmarking approach. According to [12], if a model is purely descriptive, the application of the model is seen as a single, timely assessment with no provision for improvement. A prescriptive model indicates how to approach maturity improvement in order to positively affect the value of the business/ process. The comparative model allows benchmarking in industries or regions [12]. Of the two articles analyzed, [10,11] describe a comparative application of CMMI, in which process area evaluations are compared between companies.

CONCLUSIONS

This study has enabled the main objective of this work to be achieved. When reviewing the 36 articles, it was noticed that several MMs had been applied in project management, such as OPM3, KPMMM, P2M3, P3M3, P2CMM and PMMM. However, the focus of application of these articles is on segments that show little interest in software development. We verified that this segment limits itself to using the traditional MM of SEI, the CMMI. It is believed that this occurs due to the worldwide scope that these models have, which has led to their having become firmly established as among the most frequently and widely used MMs in software engineering. However, models such as PMMM, OPM3 and KPMMM are seen to have been applied widely in several segments. Elements and features of the PMBOK guide, such as evaluation software, are embedded within the structure of the models and their application. These characteristics become so attractive that this justifies applying the model in other domains, such as in software engineering.

Finally, the main limitation of this study is the lack of an in-depth comparative approach to examining the use of MMs in project management. It would be interesting for the academic literature to identify the main similarities/ differences, as well as advantages/ disadvantages of each of the MMs developed/ applied for this domain. Therefore, for future studies, it is suggested that a comparative study that has broad coverage on MMs in project management should be conducted.

REFERENCES

[1] Y. Y. L. Helgesson, M. Host, K. Weyns. "A review of methods for evaluation of maturity models for process improvement." Journal of Software-Evolution and Process 24 (4):436-54, 2012.
[2] J. B. S. S. Santos-Neto, A. P. C. S. Costa. “Enterprise Maturity Models: a Systematic Literature Review.” Enterprise Information Systems, v. 13, p. 1-51, 2019.
[3] J. F. Blackstone Jr, J. H. Cox III, J. G. Schleier Jr. “A tutorial on project management from a theory of constraints perspective.” International Journal of Production Research, 47 (24), 7029-7046, 2009.
[4] Software Engineering Institute. “The Capability Maturity Model Guidelines of improving the Software Process”. Reading MA： Addition-Wesley, 1994.
[5] Z. Lianying, J. He; X. Zhang. “The project management maturity model and application based on PRINCE2”. Procedia Engineering, v. 29, p. 3691-3697, 2012.

[6] M. D. A. Marconi, E. M. Lakatos. “Técnicas de pesquisa”. 5ª edição. São Paulo-SP: Editora Atlas, 2002.
[7] N. J. Van Eck, L. Waltman. “Software Survey: VOSviewer, a Computer Program for Bibliometric Mapping.” Scientometrics 84 (2): 523–538, 2010.
[8] C. H. Cheng, J. R. Chang, C. Y. Kuo. “A CMMI appraisal support system based on a fuzzy quantitative benchmarks model.” Expert Systems with Applications, v.38 (4):4550–8, 2011.
[9] M. Kalpana, K. Tamizarasu, A. E. Jeyakumar. “A fuzzy logic based framework for assessing the maturity level of Indian small-scale software organizations.” Computer Systems Science and Engineering, v. 29 (2):153–67, 2014.
[10] P. Crosby. “Quality Is Free”. New York: McGraw Hill, 1979.
[11] Software Engineering Institute. “Capability Maturity Model Integration (CMMI-SE/SW/IPPD/SS)”. Version 1.1, Carnegie Mellon University, Pittsburgh, 2002.
[12] T. De Bruin, R. Freeze, U. Kaulkarni, M. Rosemann. “Understanding the Main Phases of Developing a Maturity Assessment Model.” Paper presented at the 16th Australasian Conference on Information Systems, Sydney, Australia, Nov 30–Dec 2, 2005.

Appendix A – List of papers reviewed

	Reference
Paper 1	Albrecht, J. C., and K. Spang. 2014. “Linking the benefits of project management maturity to project complexity: Insights from a multiple case study”. International Journal of Managing Projects in Business, 7(2), 285-301.
Paper 2	Abdul Rasid, S. Z., W. K. Wan Ismail, N. H. Mohammad, and C. S. Long. 2014. "Assessing Adoption of Project Management Knowledge Areas and Maturity Level: Case Study of a Public Agency in Malaysia." Journal of Management in Engineering 30 (2):264-71.
Paper 3	Bay, A. F., and Skitmore, M. 2006. Project management maturity: some results from Indonesia. Journal of Building and Construction Management, 10, 1-5.
Paper 4	Backlund, F., D. Chronéer, and E. Sundqvist. 2015. “Maturity Assessment: Towards Continuous Improvements for Project-Based Organisations?” International Journal of Managing Projects in Business, 8(2), 256-278.
Paper 5	Cooke-Davies, T. J., and A. Arzymanow. 2003. “The Maturity of Project Management in Different Industries: An Investigation into Variations between Project Management Models.” International Journal of Project Management, 21(6), 471-478.
Paper 6	Carvalho, P. V. R., C. H. S. Grecco, A. M. Souza, G. J. Huber, and J. O. Gomes. 2016. “A Fuzzy Model to Assess Disaster Risk Reduction Maturity Level Based on the Hyogo Framework for Action.” Natural Hazards 83 (1): 309–26.
Paper 7	Cheng, C. H., J. R. Chang, and C. Y. Kuo. 2011. "A CMMI appraisal support system based on a fuzzy quantitative benchmarks model." Expert Systems with Applications 38 (4):4550-8.
Paper 8	Diaz, M., and J. Sligo. 1997. "How software process improvement helped Motorola." IEEE Software 14 (5):75-81.
Paper 9	Dzazali, S., A. Sulaiman, and A. H. Zolait. 2009. “Information Security Landscape and Maturity Level: Case Study of Malaysian Public Service (MPS) Organizations.” Government Information Quarterly 26 (4): 584–93
Paper 10	Fuad , B. A., M. Skitmore, and M. Skitmore. 2005. “Project Management Maturity: Some Results From Indonesia.” Indonesia. Journal of Building and Construction Management 10 (September 2005): 1–5.
Paper 11	Grant, K. P., and J. S. Pennypacker. 2006. "Project management maturity: An assessment of project management capabilities among and between selected industries." IEEE Transactions on Engineering Management 53 (1):59-68.

Paper 12	Ghoddousi, P., Z. Amini, and M. Hosseini. 2011. "A Survey on the Maturity State of / Ran / an Grade One Construction Companies Utilizing OPM3 Maturity Model" 6 (1).
Paper 13	Gilchrist, J. M. 1992. "Project evaluation using the SEI method." Software Quality Journal 1 (1):37-44.
Paper 14	Kalpana, A. M., K. Tamizarasu, and A. E. Jeyakumar. 2014. "A fuzzy logic based framework for assessing the maturity level of Indian small-scale software organizations." Computer Systems Science and Engineering 29 (2):153-67.
Paper 15	Kuvaja, P., and A. Bicego. 1994. "BOOTSTRAP - A European assessment methodology." Software Quality Journal 3 (3):117-27.
Paper 16	Lianying, Z., Jing, H., & Xinxing, Z. (2012). The project management maturity model and application based on PRINCE2. Procedia Engineering, 29, 3691-3697.
Paper 17	Lowe, D. E., and G. M. Cox. 1996. "Implementing the capability maturity model for software development." Hewlett-Packard Journal 47 (4):6-14.
Paper 18	Li, E. Y., H. G. Chen, and T. S. Lee. 2002. "Software process management of top companies in Taiwan: a comparative study." Total Quality Management 13 (5):701-13.
Paper 19	Mittermaier, H.K., and H. Steyn. 2009. "Project Management Maturity: An Assessment of Maturity for Developing Pilot Plants." The South African Journal of Industrial Engineering 20: 95–107.
Paper 20	Mcgrew, J. F., J. G. Bilotta, and J.M. Deeney. 1999. "Software Team Formation and Decay : Extending the Standard Model for Small Groups Extending the Standard Model" 30 (2): 209–34.
Paper 21	Macchi, M., and L. Fumagalli. 2013. "A Maintenance Maturity Assessment Method for the Manufacturing Industry." Journal of Quality in Maintenance Engineering, 19(3), 295-315.
Paper 22	McCuen, T. L., P. C. Suermann, and M. J. Krogulecki. 2012. "Evaluating Award-Winning BIM Projects Using the National Building Information Model Standard Capability Maturity Model." Journal of Management in Engineering 28 (2):224-30.
Paper 23	Neves, L. A. Leite das, L. E. Nicolini do P. Nunes, V. A. Corrêa, and M. C. Rezende. 2013. "Application of the Prado - Project Management Maturity Model at a Rand Institution of the Brazilian Federal Government." Journal of Aerospace Technology and Management 5 (4): 459–65.
Paper 24	Pretorius, S., H. Steyn and J. C. Jordaan. 2012. "Project management maturity and project management success in the engineering and construction industries in Southern Africa". South African Journal of Industrial Engineering, 23(3), 1-12.
Paper 25	Phan, D. D. 2001. "Software quality and management - How the world's most powerful software makers do it." Information Systems Management 18 (1):56-67.
Paper 26	Puus, U., and T. Mets. 2010. "Software development maturity evaluation: six cases from Estonian SMEs." Baltic Journal of Management 5 (3):422-43.
Paper 27	Rendon, R. G. 2015. "Benchmarking contract management process maturity: a case study of the US Navy". Benchmarking: An International Journal, 22(7), 1481-1508.
Paper 28	Silva, D., A. Tereso, G. Fernandes, and J. Pinto. 2014. "OPM3® Portugal Project: Analysis of Preliminary Results." Procedia Technology, 16, 1027-1036.
Paper 29	Siew, R. Y., M. C. Balatbat, and D. G. Carmichael. 2016. "Measuring Project Sustainability Maturity Level-a Fuzzy-Based Approach." International Journal of Sustainable Development, 19(1), 76-100.
Paper 30	Seong Leem, C., and Y. K. Yoon. 2004. "A maturity model and an evaluation system of software customer satisfaction: the case of software companies in Korea." Industrial Management and Data Systems 104 (3-4):347-54.
Paper 31	Tiku, S., M. Azarian, and M. Pecht. 2007. "Using a Reliability Capability Maturity Model to Benchmark Electronics Companies." International Journal of Quality and Reliability Management 24 (5): 547–63.
Paper 32	Vander Velden, M. J., J. Vreke, B. vanderWal, and A. Symons. 1996. "Experiences with the Capability Maturity Model in a research environment." Software Quality Journal 5 (2):87-95.
Paper 33	Yazici, H. J. 2009. "The role of project management maturity and organizational culture in perceived performance." Project Management Journal, 40(3), 14-33.
Paper 34	Young, M., R .Young, and J. Romero Zapata. 2014. "Project, Programme and Portfolio Maturity: A Case Study of Australian Federal Government." International Journal of Managing Projects in Business, 7(2), 215-230.
Paper 35	Wetering, R. V., R. Batenburg, M. Oudkerk, P. Van Ooijen, S. Brinkkemper, and W. Scheper. 2011. "A Situational Alignment Framework for PACS." Journal of Digital Imaging 24 (6): 979–92.
Paper 36	Zhong, L. N., D. Leung, R. Law, B. H. Wu, and J. Shao. 2014. "An Application of the Capability Maturity Model for Evaluating Attraction Websites in Mainland China." International Journal of Tourism Research 16 (5):429-40.

Comparison of Normalization Techniques for Skewed Data Sets

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

ABSTRACT

With the fast growing of data-rich systems and big data, dealing with complex decision problems with skewed input data sets and respective outliers is unavoidable. Generally, data skewness refers to a non-uniform distribution in a dataset, i.e. a dataset which contains asymmetries and/or outliers. Normalization is the first step of most multi-criteria decision making (MCDM) problems to obtain dimensionless data, from heterogeneous input data sets, that enable aggregation of criteria and thereby ranking of alternatives. Therefore, when in presence of skewness in criterion datasets, finding a suitable normalization technique is of utmost importance. As such, in this work, we compare four normalization techniques (Max, Max-Min, and Vector, Target-based) on criteria datasets, which contain outliers to analyze their results for MCDM problems. A numerical example displays the behavior of the four normalization techniques and Step B (Pearson correlation) of an (ongoing) evaluation assessment framework is used to recommend the best normalization technique for this type of criteria In the extended version all steps, of the on-going evaluation assessment frame work, will be used to ensure the choice is appropriate.

Keywords: Normalization, Data Set, MCDM, Decision Making, Outliers, Target value, skewed data.

Multicriteria Decision Making Approaches

Dynamic-R: a new "convincing" multiple criteria method for rating problem statements

Oussama Raboun, Tsoukias Alexis
Univ. Paris-Dauphine, PSL Research University, CNRS, UMR [7243], LAMSADE, 75016 PARIS, FRANCE,
oussama.raboun@dauphine.eu,
tsoukias@lamsade.dauphine.fr

Eric Chojnacki
Institut de Radioprotection et de Sureté Nucléaire (IRSN), CADARACHE, FRANCE
eric.chojnacki@irsn.fr

ABSTRACT

In this paper, we propose Dynamic-R method, a new decision aiding procedure to deal with multicriteria rating problems. A multicriteria rating problem consists on partitioning a set of objects, assessed under several dimensions called criteria, into predefined ordered equivalence classes, called categories, identified by rates. Several rating methods were developed using the majority rule. These methods present many disadvantages leading potentially to an unconvincing rating. In this work, we introduce a dynamic rating procedure aiming at providing a "convincing" rating (stable under criticisms) over a set of studied objects. It is called dynamic, since rated objects will be used to characterize the categories in the next steps. The developed rating procedure is based on the aggregation of positive and negative reasons respectively supporting and opposing to a rating.

Keywords: Multicriteria Decision Aiding, Rating Problem Statements, Decision Support Systems, Algorithmic Decision Theory

Find and cite the full work as:
Raboun, O., Chojnacki, E., Tsoukias, A., 2019. Dynamic-R: a new "convincing" multiple criteria method for rating problem statements, in: Freitas, P.S.A., Dargam, F., Moreno, J.M. (Eds.), Decision Support Systems IX: DSS Developments & Future Trends, Lecture Notes in Business Information Processing. Springer International Publishing, Cham, pp. 129-141.

Intelligent Multicriteria Decision Support System for a Periodic Prediction

Sarra Bouzayane, Ines Saad
MIS Laboratory, University of Picardie Jules Verne, Amiens 80039, France
Amiens Business School, Amiens 80039, France
{sarra.bouzayane,ines.saad}@u-picardie.fr

ABSTRACT

This paper proposes an intelligent decision support system for the Incremental Periodic Prediction of the decision class to which an action is likely to belong. This method is based on three phases. The first consists of three steps: the construction of a family of criteria for the characterization of actions; the construction of a representative learning set for each of the decision classes; and the construction of a decision table. The second phase is based on the DRSA-Incremental algorithm that we propose for the inference and the updating of the set of decision rules following the sequential increment of the learning set. The third phase is meant to classify the "Potential Actions" in one of the predefined decision classes using the set of inferred decision rules. Our method is based on the DRSA (Dominance-based Rough Set Approach) which is a supervised learning technique permitting to extract the preferences of decision makers for the actions categorization. Applied in the context of MOOCs (Massive Open Online Courses) for the categorization of learners' profiles, our approach proved a satisfactory classification quality that reaches 0.66.

Keywords: Intelligent Decision Support System, Multicriteria decision making, Incremental prediction, MOOC

Find and cite the full work as:
Bouzayane, B., Saad, I., 2019. Intelligent Multicriteria Decision Support System for a Periodic Prediction, in: Freitas, P.S.A., Dargam, F., Moreno, J.M. (Eds.), Decision Support Systems IX: DSS Developments & Future Trends, Lecture Notes in Business Information Processing. Springer International Publishing, Cham, pp. 90-102.

MCDM model to select a location for a natural gas pressure reducing station

Naylil Liria Baldin de Lacerda[1], Carolina Lino Martins [1], João Batista Sarmento dos Santos Neto[1]

[1]Universidade Federal de Mato Grosso do Sul
Cidade Universitária, S/N, Campo Grande, Mato Grosso do Sul, Brazil
naylilacerda@gmail.com, carol_tcch@hotmail.com, joao.sarmento@ufms.br
web-page: https://faeng.ufms.br/graduacao/bacharelado/engenharia-de-producao/

ABSTRACT

This paper aims to propose a multicriteria decision model to solve a location selection problem for a pressure reducing station in a natural gas distribution company. The case study under study is a Brazilian natural gas distribution company, since the chosen company is planning to expand its operations. The preference modeling was conducted through two approaches: PROMETHEE II and the Flexible Interactive Tradeoff (FITradeoff) method, once they are intuitive and straightforward to apply. The results obtained from both methods were similar, which proved the robustness of the model and that the DM evidenced consistency in its preferences. Also, the FITradeoff method demonstrated to be more intuitive to apply, since a smaller effort is required from the DM and this is because the procedure does not require complete information from the DM in the weight elicitation process.

Keywords: Natural Gas Station Location, Pressure Reducing Station, MCDM /A, PROMETHEE II, FITradeoff.

INTRODUCTION

The distribution of natural gas to the several consumer segments must be carried out uninterruptedly, with specific pressure and flow for each customer [1]. In Brazil, for instance, the expansion of gas pipelines is a responsibility of the distribution companies, and it is planned to increasingly stimulate natural gas growth, competitiveness in the sector, as well as meet the diverse demands of its stakeholders from the public and private initiative.

These organizations have come across with several difficulties in delivering satisfactory results, an efficient strategic and operational planning process that addresses the varied interests of all their stakeholders. Among these difficulties, it is possible to mention the decision making in the selection of strategic locations for the implementation of pressure reducing stations (PRS), which are used to control the outlet pressure of the natural gas in the distribution operations in order to maintain the necessary flow to meet customer's demand [1].

Such a decision must consider different financial criteria, technical operating standards, and environmental policies, besides being aligned with the company's expansion plan. In this context, Multicriteria Decision Making area is applied to support the decision making process, when it is not possible to represent all the objectives in only one project [2]. They consist in analyzing the possible implementation of an action, assessing its positive and negative characteristics comparatively, side by side. According to the level of importance of the criteria, they are evaluated in an eliminatory way, until one action is considered better [3].

Thus, this study aims to propose a multicriteria decision model to solve a location selection problem for a pressure reducing station in a natural gas distribution company. To do so, an analysis was performed in a Brazilian natural gas distribution company, since the chosen company is planning to expand its operations and must carry out studies of this nature.

To select the location for the pressure reduction station, the preference modeling was conducted through a non-compensatory approach, aided by the Preference Ranking Organization Method for Enrichment Evaluation II (PROMETHEE II) [4], an outranking method, and also a compensatory approach, assisted by the Flexible and Interactive Tradeoff method (FITradeoff), a value function method [5]. These methods were chosen because they are simple and intuitive to apply, allows the achievement of a complete ranking of alternatives, and the elicitation of them can be conducted with the decision maker (DM) by means of Decision Support Systems (DSS). Besides, it is essential to compare the results when different kinds of approaches are considered.

METHODOLOGICAL PROCEDURES

The Multiple Criteria Decision Making (MCDM) methods used were PROMETHEE II and FITradeoff. For the PROMETHEE II method, the decision making process was performed through an outranking relation [3]. The first step of the PROMETHEE II method is to elucidate the absolute weights of the criteria, the weights of this method must clearly show the preferences of the decision maker; and the definition of the type of preference function. Thus, the PROMETHEE II method performs the analysis through a pairwise comparison for each criterion, to obtain the preference relations between two alternatives. The analysis between alternatives *a* and *b* for a criterion *j* is given by a preference function P_j (a, b), which is expressed by the mathematical representation that compares one alternative with the other. In the strict preference function, called *usual* function, we have P_j (a, b) = 1 when g_i (a)> g_i (b), and in its opposite one has, P_j (a, b) = 0 [6].

The preference indices π (a, b) are calculated considering the weight w_j of each criterion *j*. Then, the outranking relation is carried out, in which the positive preference flows Φ + (a) and negative Φ - (a) are calculated. Finally, a complete ranking of the alternatives is elaborated and the alternative that presents the highest absolute value of Φ (a), is the solution to the problem [6]. Figure 1, adapted from [4], describes the procedures to apply PROMETHEE II.

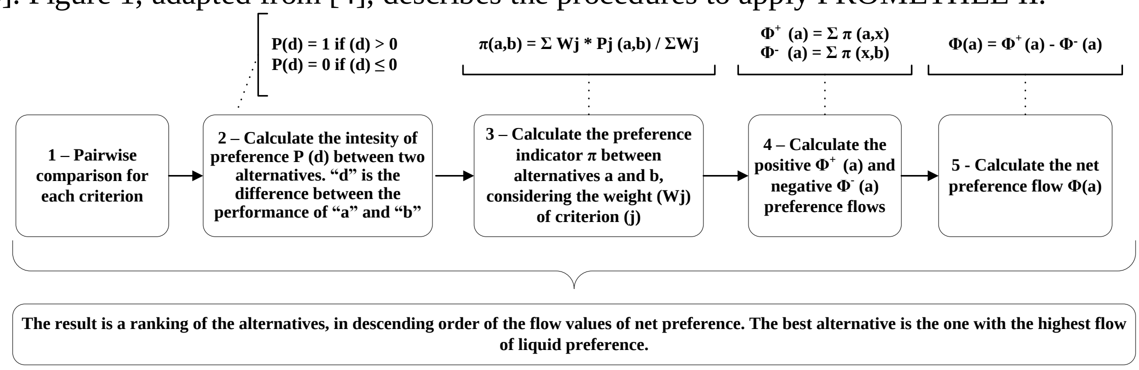


Figure 1 – PROMETHEE II method

The FITradoff method is based on tradeoffs, used for the elicitation of weights (k_i) of criteria in additive models [5]. [5] proposed the FITradoff method with a strong axiomatic tradeoff structure, in a simple process to determine the performances (*v*) of the alternatives (*a*) in the criteria (*i*) when using partial information of the decision maker (DM) [5], allowing the decrease of inconsistencies in the model. The FITradeoff process is summarized in Figure 2, which is based on [5].

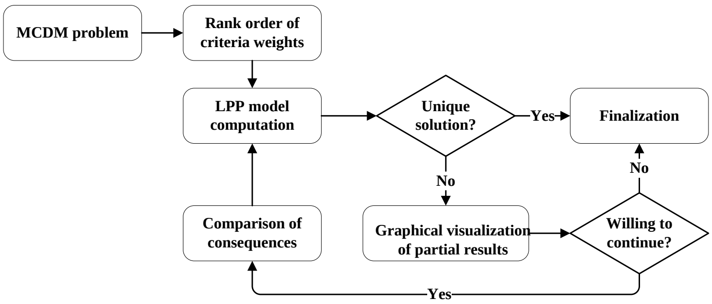


Figure 2 – FITradoff method

When considering a multicriteria problem with *m* alternatives and *n* criteria, the first step of the FITradeoff is to order the weights for the criteria according to the preferences of the DM, according to (1.0), where k_1 is the weight for the most preferable criterion and k_n the least preferable [5].

$$k1 > k2 > k3 > \cdots > kn \quad (1.0)$$

Thus, all alternatives are verified by solving a linear programming problem, and an alternative is possibly optimal if (2.0) is satisfied [5].

$$\sum_{i=1}^{n} ki\ vi\ (xij) \geq \sum_{i=1}^{n} ki\ vi\ (xiz) \quad \forall z = 1, \ldots, m, z \neq j \quad (2.0)$$

The optimal value of an alternative is verified through the maximization of the value a_j, according to (3.0), in a linear programming model, where the decision variable is the weight k_i, submitted to (1.0) and (2.0) [5].

$$Max \sum_{i=1}^{n} ki\ vi\ (xij) \quad (3.0)$$

Thus, the process is performed to find the solution of all *m* alternatives until only an alternative a_j is found. If the solution for the alternatives is not found, the method moves on to the second step, in which the decision maker must compare two possible consequences, considering the tradeoffs [5].

In this step, preference relations and partial information are used in inequations (4.0) and (5.0), which will be part of a new linear programming problem as new constraints, to find the solution to the problema [5].

$$ki\ vi(x_i') \geq k_{i+1} \quad (4.0)$$

$$ki\ vi(x_i'') \geq k_{i+1} \quad (5.0)$$

FITradeoff is an interactive method and the second step must be performed until a single solution is found or until the decision maker stops providing additional information [5].

RESULTS AND DISCUSSION

The area where the pressure reduction station is to be implemented is located in a potential region for the expansion of the natural gas distribution network, since new urban roads will be opened, which will allow the creation of integration axes and local development, together with the expansion generated by housing programs from the local government, providing the distribution of natural gas to the commercial and residential segments.

Since it is a complex problem, in this decision process occurs the situation in which there is the influence of several decision making actors; however, only one individual is directly responsible for the decision in question [3]. Therefore, the DM was defined as the director of the technical and commercial area; and as specialists, the managers, who are responsible for relevant information to the problem.

All actors in the decision making process were consulted and established basic specifications for choosing a location that includes a pressure reduction station, such as the location should be easily operable, with the possibility of interconnecting the new networks of carbon steel (CS) and high density polyethylene (HDPE) to the existing network, not susceptible to flooding, and with the possibility of interconnection of water and electric energy; regarding environmental policy, the location must be licensed to build and maintain a pressure reducing station and must meet legal constraints, such as neighborhood area, zoning and building codes. Taking this into account, the experts surveyed the places that meet the basic specifications and that are close to the region of interest, as well as their characteristics. In addition, they carried out studies on the resources needed to implement the pressure reduction station at each of the potential sites.

After defining the basic specifications, the criteria were identified as the area, the extension of high density polyethylene pipes and carbon steel, the impact on the neighborhood, the location cost and network deployment cost. The size of the area was considered since the site should support a pressure reduction station; the extension of HDPE and CS pipes were chosen as criteria because they influence the resources needed for the deployment, so the location should be as close as possible to the existing network; the impact on the neighborhood was considered due to the fact that there are environmental parameters that must be met; location and network deployment costs are important financial criteria that should be minimized.

The problem of finding a location that meets the criteria for the construction of the secondary pressure reducing station has enabled the development of a multicriteria decision model that is meant to support the decision making process. With the set of alternatives and the criteria, the decision matrix was identified, as observed in Table 1.

Table 1 – Decision matrix

Alternatives/ Criteria	Area (m²)	Carbon Steel pipe (m)	High density Polyethylene pipe (m)	Neighbor Impact	Location cost (R$)	Pipe cost (R$)
L1	451	280	4.210	Baixo	R$ -	R$ 1.936.000,00
L2	720	-	3.440	Alto	R$ 1.000.000,00	R$ 2.376.000,00
L3	1.436	1.540	2.620	Alto	R$ 660.000,00	R$ 3.094.000,00
L4	750	700	2.730	Alto	R$ -	R$ 1.722.000,00
L5	765	2.240	1.200	Alto	R$ 385.000,00	R$ 2.881.000,00
L6	900	1.730	1.920	Baixo	R$ 330.000,00	R$ 2.655.000,00
L7	10.000	1.450	2.300	Alto	R$ -	R$ 2.225.000,00
L8	1.549	1.680	2.750	Médio	R$ 400.000,00	R$ 3.012.000,00
Weight	2	5	6	8	9	9
Objective	Maximize	Minimize	Minimize	Minimize	Minimize	Minimize

The evaluation of the alternatives, as well as the data for analyzes and comparisons, were obtained through the Visual PROMETHEE software and the FITradeoff DSS. In the PROMETHEE II method, the criteria weights were normalized by the procedure of division by the sum of the values so that the alternatives performed a pairwise comparison [6]. And in this paper, it was considered the usual function in where no particular parameter, such as indifference and preference thresholds, has to be defined and it gives the opportunity to the DM to use the criterion in its usual sense when required [3].

In the FITradeoff method the performance values of each alternative in each criterion are normalized in a ratio 0–1 scale. The FITradeoff DSS presented the optimal solution in the first step of the procedure, that is, when the DM ordered the criteria according to their preferences, showing that the method requires less cognitive effort from the DM during the elicitation

process and this is because the method works with partial information about the DM's preferences.
Among the eight alternatives that could have presented themselves as the best compromise solution for the multicriteria problem, in both methods, the ranking presented the alternative L4 as the best location option to implement the pressure reducing station.
PROMETHEE II granted a descending complete pre-order of the alternatives, according to the net flow Φ (a), evidenced in Table 2. As well as enabled the evaluation of each criterion for each alternative through Table 3, in which the alternative L4 presented a good performance for all criteria, mainly in the financial parameters. The cost of network deployment presented itself as the criterion with the best performance, since the combination of CS and HDPE network extension provided the lowest cost for the company.

Table 2 – PROMETHEE II results

Classification das Alternatives	Φ(a)	Φ+ (a)	Φ- (a)
L4	0,4945	0,6557	0,1612
L7	0,4872	0,652	0,1648
L1	0,0403	0,4725	0,4322
L2	-0,0256	0,4286	0,4542
L5	-0,1282	0,3773	0,5055
L6	-0,1502	0,4103	0,5604
L3	-0,2821	0,3004	0,5824
L8	-0,4359	0,2821	0,7179

Table 3 – Preference flows

Alternatives/ Criteria	Area (m^2)	Carbon Steel pipe (m)	High density Polyethylene pipe (m)	Neighbor Impact	Location cost (R$)	Pipe cost (R$)
L1	-1	0,7143	-1	-0,8571	0,7143	0,7143
L2	-0,7143	1	-0,7143	0,4286	-1	0,1429
L3	0,4286	-0,1429	0,1429	0,4286	-0,7143	-1
L4	-0,4286	0,4286	-0,1429	0,4286	0,7143	1
L5	-0,1429	-1	1	0,4286	-0,1429	-0,4286
L6	0,1429	-0,7143	0,7143	-0,8571	0,1429	-0,1429
L7	1	0,1429	0,4286	0,4286	0,7143	0,4286
L8	0,7143	-0,4286	-0,4286	-0,4286	-0,4286	-0,7143

In the FITradeoff method, the elicitation of the weights was performed according to DM preferences. Table 4 presents the values of the scale constants that maximize the global value of the alternative L4. When analyzing the results provided by FITradeoff, it was observed that the distribution patterns of the weights did not have significant variations, since ranking the criteria does not provide much information for these standards.

Table 4 – FITradeoff results

ALTERNATIVE	K(C5)	K(C6)	K(C4)	K(C3	K(C2)	K(C1)	Maximum Value
L4	0,5161	0,2581	0,129	0,0645	0,0323	0	0,8849

The tool "Visual Stability Intervals" from the Visual PROMETHEE software was used to perform a sensitivity analysis, to calculate the upper and lower boundaries of the results, i.e., the limits of the weight in which every criterion keeps the ranking unchanged when the other weights are kept at their initial levels. The neighbor impact criterion is the most sensitive, meaning that a small variation in its value has a bigger impact on the results when compared with other criteria. On the other hand, the pipe cost criterion is the least sensitive, which means that a big variation in its value is needed to influence the results. The sensitivity analysis results help the DM to see the impact of each criterion in the model.

The FITradeoff software also allows the performance of a sensitivity analysis. The process was carried out using ten thousand simulations with a variation of ± 10% in the weights of all the criteria. The sensitivity analysis of the FITradeoff presented the inclusion of the L7 alternative in the original subset of potentially optimal alternatives, with 24.34% of occurrence.

CONCLUSIONS

In this study, the application of two multicriteria approaches, PROMETHEE II and FITradeoff, was presented to assess the decision making process of selecting a location for the implementation of a natural gas pressure reducing station.

The application was the main difference between the two methods. In PROMETHEE II, the DM evaluates and ponders the absolute weights of the criteria individually, considering the interests of the stakeholders. In the FITradeoff method, the elicitation of the weights of the criteria occurs through comparative questions, which allows a better understanding and evaluation of the DM regarding the relevance of the criteria that will be used to define a solution that is close to reality.

The results from both methods were similar, which allows inferring that the proposed multicriteria model was robust, satisfactory and that the DM showed consistency in its preferences. It was also noted that FITradeoff requires less cognitive effort from the DM during the elicitation process and this is because the method works with partial information about the DM's preferences.

Decision problems like the one presented in this study are common in other natural gas distribution companies. Hence, this study can be replicated or used as a basis by other organizations, so that decisions are not empirically made.

REFERENCES

[1] M. H. Mohammed, "Design of an ideal gas regulating and metering station for gas supply to A 50 MW power plant.", 2009

[2] F. Burstein, C. W. Holsapple (Eds.), "Handbook on decision support systems 2: variations." Springer Science & Business Media, 2008.

[3] T. de Almeida, C. A. V. Cavalcante, M. H. Alencar, R. J. P. Ferreira, A. T. de Almeida-Filho, T. V. Garcez, "Multicriteria and Multiobjective Models for Risk, Reliability and Maintenance Decision Analysis." Springer, Cham, 2015.

[4] J. P. Brans, P. Vincke, B. Mareschal, "How to select and how to rank projects: The PROMETHEE method." European journal of operational research, 24(2), 228-238, 1986.

[5] T. de Almeida, J. A. de Almeida, A. P. C. S. Costa, A. T. de Almeida-Filho, "A new method for elicitation of criteria weights in additive models: Flexible and interactive tradeoff." European Journal of Operational Research, 250(1), 179-191, 2016.

[6] J. P. Brans, B. Mareschal, Promethee Methods: Multiple-Criteria Decision Analysis. State of the Art Surveys (Chapter 5). "Springer International Series in Operations Research and Management Science". New York: Springer, 2005.

Multiple criteria decision making model to assess organizational climate in healthcare institutions

Kassia Tonheiro Rodrigues[1,2], Carolina Lino Martins[1], João Batista Sarmento dos Santos Neto[1], Diego Rorato Fogaça [1]

[1]Universidade Federal de Mato Grosso do Sul
Cidade Universitária, S/N, Campo Grande, Mato Grosso do Sul, Brazil
carol_tcch@hotmail.com, joao.sarmento@ufms.br, diego.fogaca@ufms.br
web-page: https://faeng.ufms.br/graduacao/bacharelado/engenharia-de-producao/

[2]Universidade Federal de Santa Catarina
R. Eng. Agronômico Andrei Cristian Ferreira, S/N - Trindade, Florianópolis, state of Santa Catarina, Brazil
kassia.tonheiro@gmail.com

ABSTRACT

This work aims to propose a model to assess organizational climate in healthcare institutions. To do so, a study was made in a Clinical Analysis Laboratory Unit of a Brazilian public health organization using Multicriteria Decision Making /Aid (MCDM /A). The multicriteria decision method applied was the ELECTRE TRI (Elimination and Choice Translating Reality), once it was an outranking problem. A review of the literature was carried out to establish the model criteria, including internal and external factors that can influence the organizational climate. To elicit the weights of the model, an aggregation of experts knowledge was performed, and the model was conducted by means of a Decision Support System (DSS). As a result, the model allowed to sort the organizational climate as favorable in an optimistic scenario and less favorable in the pessimistic scenario. The main contribution of this work is the validation of MCDM/A method in assessing organizational climate in healthcare institutions.

Keywords: MCDM/A, Organizational Climate, ELECTRE TRI, Healthcare.

INTRODUCTION

Healthcare institutions provide several services for the population. They are inserted in a scenario of changes concerning the patients' demands, increase the life expectancy of people and demographic variations. All these issues require complex and constantly evolving management that seeks to add and create value to services offered [1]. Besides that, it’s important to emphasize that there is a difference between managing a public and a private hospital, because it is necessary to optimize resources when dealing with a public organization, including those that have a scarce supply and cannot be offered financial advantages to the servers [2].

In these circumstances, the concept of organizational climate is highlighted, which can be considered as a set of multidimensional perceptions of the individuals belonging to such organization. These perceptions develop with the meaning taken by each one in the

organizational context according to their pre-existing values [3, 4].
This means that the organizational climate can have a direct impact on employees' performance and well-being. Thus, the research of organizational climate has been a tool widely used by organizational psychology, in organizations of different branches. This analysis allows the organization to know what the individual's perception of the environment [4, 5].
Therefore, this paper aims to propose a model to assess organizational climate in healthcare institutions. The preference modeling was conducted through a non-compensatory approach, aided by the Elimination and Choice Translating Reality (ELECTRE TRI), an outranking method [6]. MCDA methods provide an essential and efficient way to support the decision-making process, considering multiple criteria, and incorporating the decision maker (DM) preference structure. ELECTRE TRI is one of the most used methods for multicriteria sorting problematic [7]. In this study, the application of ELECTRE TRI was assisted by the Decision Support System (DSS) *MDCA - Ulaval*.

METHODOLOGY

The ELECTRE TRI method was used to sort the organizational climate of a Clinical Analysis Laboratory Unit of a Brazilian public health organization, since the solution of the problem brings to the DM a better understanding regarding the perception of the employees, which helps the decision-making process to improve organizational climate. ELECTRE TRI seeks to assign an alternative to any specific category based on a pairwise comparison of its evaluation with a particular set of criteria and categories limiting profiles [6].
The model proposed by [8] was chosen to define the criteria of the model, since it has a broad approach when analyzing internal and external factors regarding organizational climate. Also, it considers some factors of previous studies and additional factors that emerged with globalization, technological innovations, diffusion of information and increased competition among companies. The set of criteria are presented in Table 1.
To elicit the weights of the model an aggregation of experts knowledge was performed. Hence, six professors from an engineering program of a public university in Brazil were selected as experts to define their preferences thorough a questionnaire, where 0 would be the criterion of minor relevance and 100, the major.

Table 2: Set of criteria to assess organizational climate

INTERNAL FACTORS	EXTERNAL FACTORS
Structure and rules	Family living
Relationship and cooperation	Holidays and leisure
Identity and pride	Physical and mental health
Leadership and support	Family financial situation
Consideration, prestige and tolerance	Politics and economics
Opportunity for growth	Public security
Organizational culture	Social life
Stability in employment	Performance of your sport team
Transport (between home and work)	
Sociocultural level	

In this research, internal factors and external factors were considered as independent alternatives. A three-points scale questionnaire was constructed, where 1 corresponds to non-

favorable climate, 2 less favorable climate and 3 favorable climate. Thirty-four questionnaires were distributed in the laboratory, in which twenty-three were answered. The mode of the answers was aggregated to define the performance of each criterion.
For the ELECTRE TRI application, the preference thresholds (*p*), indifference (*q*) and the veto threshold (*v*) must be determined. Therefore, in this study we chose to assign the zero value to the thresholds, which brings a notion of true criterion rather than pseudo-criterion, as it is considered in ELECTRE TRI. This means that a zone of hesitation, represented by the weak preference, is avoided [10, 11].
The method is based on two steps [6]: construction of an outranking relation '*S*' that is characterized by comparing the alternatives to the categories profiles, and; exploitation of the relation *S* to assign each alternative to a specific category. For further information on ELECTRE TRI procedures, see [6, 9].
To run the *MDCA - Ulaval* it was necessary to establish the cutting level (λ), where $0,5 \leq \lambda \leq 1$, and the limiting profiles for the classes. The profiles used for applying the method are presented in Table 2.

Table 2: Classes and Profiles for the ELECTRE TRI Method

NAME	DESCRIPTION	NUMBER
Classe 3	Favorable	
Profile 2		2
Classe 2	Less favorable	
Profile 1		1
Classe 1	Not favorable	

The decision matrix data, weights, preference and indifference thresholds, veto threshold, cutting level, profiles, and classes were entered into the software. This step was performed twice: first to sort the organizational climate evaluated on the approach of external factors and secondly to evaluate internal factors.

RESULTS AND DISCUSSION

The mode of the answers regarding each criterion was identified and are described in Table 3.

Table 3: Criteria from the model and mode of the answers

	CRITERIA	MODE
INTERNAL FACTORS	Identity and pride	3
	Opportunity for growth	3
	Leadership and support	3
	Structure and rules	2
	Consideration, prestige and tolerance	3
	Stability in employment	3
	Sociocultural level	3
	Transport (between home and work)	1
	Relationship and cooperation	3
	Organizational culture	2
EXTERNAL	Family financial situation	3
	Family living	3

	Social Life	3
	Physical and mental health	2
	Performance of your sport team	3
	Holidays and leisure	3
	Public security	3
	Politics and economics	1

The preference (*p*) and indifference (*q*) thresholds were considered zero. Also, the veto value was equal to 1 (v = 1) and the cutting level equal to 0.75 ($\lambda = 0.75$). To define the weights, an average of the values attributed by the experts was made. The weights were normalized by dividing each of them by the total summation, so the sum of the weights remained equal to one, as shown in Table 4.

Each alternative was evaluated separately. In both cases, internal and external factors, the climate was sorted as less favorable in the pessimistic procedure, while in the optimistic view, as favorable. When in the pessimistic and optimistic procedures an alternative is assigned in the same category, the method performs the assignment in the category identified. However, when the procedure assigns into different categories, it is characterized as incomparable. The pessimistic procedure is considered as the most rigorous [10].

To analyze the robustness of the model, a two-stage sensitivity analysis was performed. The first step consisted of varying preference (*p*) and indifference (*q*) thresholds. In both alternatives, five tests were performed, varying the thresholds from 0.1 until 0.5, and in none of them, there was a change in the results. In the second step, the variations were performed by changing the cutting level (λ). Five tests were also performed in the assignment of each alternative, and the cutting level was varied up to $\lambda = 1$. In this process, there was no change in the results, which demonstrates the effectiveness of the proposed model and the use of the method.

Table 4: Criteria, average of the weights and standard weights

	CRITERIA	AVERAGE	STANDART WEIGHTS
INTERNAL FACTORS	Structure and rules	82.50	0.08
	Relationship and cooperation	80.83	0.08
	Identity and pride	73.83	0.07
	Leadership and support	88.67	0,08
	Consideration, prestige and tolerance	73.83	0.07
	Opportunity for growth	79.17	0.07
	Organizational culture	92.67	0.09
	Stability in employment	69.17	0.06
	Transport (between home and work)	40.00	0.04
	Sociocultural level	41.67	0.04
EXTERNAL FACTORS	Family living	54.17	0.05
	Holidays and leisure	53.17	0.05
	Physical and mental health	78.17	0.07
	Family financial situation	58.00	0.05

	Politics and economics	29.17	0.03
	Public security	28.50	0.03
	Social life	40.67	0.04
	Performance of your sport team	3.50	0.003
	Total	1067.67	1.00

CONCLUSIONS

This study presented an MCDA methodology to evaluate organizational climate in a healthcare institution. A Clinical Analysis Laboratory Unit of a Brazilian public health organization was selected to propose the model using the ELECTRE TRI method.
The results were satisfactory, considering that a questionnaire was used to collect data; the model allowed the evaluation of the organizational climate and the robustness of the results were proved with the sensitivity analysis. Also, for both internal factors and external factors, in a pessimistic scenario the climate was sorted as less favorable and in an optimistic scenario, as favorable, showing consistency in outranking.
As a contribution, this work allows the manager a vision of the employees' perceptions regarding the organizational climate and, in addition, allows identifying which factors should be prioritized in the implementation of an action plan. Besides that, the research has relevance in showing the effectiveness of the application of an MCDA methodology in the evaluation of the organizational climate, considering that in the literature, there are several questionnaires and models of research of organizational climate, which are not always accompanied by procedures for evaluation.
As limitations of the study, it can be highlighted the difficulty of establishing values for the parameters of the model, such as thresholds of preference and indifference. Finally, it can be stated that the proposed model broadens the vision of the tools that can be used in healthcare management. For future work, it is suggested to analyze the same database using other MCDA methods in order to compare the results and the efficiency of the proposed method.

REFERENCES

[1] Chow, C. W., Ganulin, D., Haddad, K., & Williamson, J. (1998). The balanced scorecard: a potent tool for energizing and focusing healthcare organization management. Journal of Healthcare Management, 43(3), 263.

[2] Clarkson, K. W. (1972). Some implications of property rights in hospital management. The Journal of Law and Economics, 15(2), 363-384.

[3] James, L. A., & James, L. R. (1989). Integrating work environment perceptions: Explorations into the measurement of meaning. Journal of applied psychology, 74(5), 739.

[4] James, L. R., James, L. A., & Ashe, D. K. (1990). The meaning of organizations: The role of cognition and values. Organizational climate and culture, 40, 84.

[5] James, L. R., & Jones, A. P. (1974). Organizational climate: A review of theory and research. Psychological bulletin, 81(12), 1096.

[6] Mousseau, V.; Slowinski, R.; Zielniewicz, P. A user-oriented implementation of the ELECTRE-TRI method integrating preference elicitation support. Computers &

operations research, v. 27, n. 7-8, p. 757-777, 2000.
[7] T. de Almeida, C. A. V. Cavalcante, M. H. Alencar, R. J. P. Ferreira, A. T. de Almeida-Filho, T. V. Garcez, "Multicriteria and Multiobjective Models for Risk, Reliability and Maintenance Decision Analysis." Springer, Cham, 2015.
[8] Bispo, C. A. F. A new model of organizational climate research. Production, v. 16, n. 2, p. 258-273, 2006.
[9] Silva, L. C. E.; Costa, A. P. C. S. IT project investments: an analysis based on a sort and rank problem. International Journal of Information Technology & Decision Making, v. 13, n. 04, p. 699-719, 2014.
[10] Miranda, C. M. G. D.; De Almeida, A. T. Postgraduate evaluation through ELECTRE TRI method - the case of III engineering area of Capes. Production, v. 13, n. 3, p. 101, 2003.
[11] Rogers, M., & Bruen, M. (1998). Choosing realistic values of indifference, preference and veto thresholds for use with environmental criteria within ELECTRE. European Journal of Operational Research, 107(3), 542-551.

Business Processes & Business Intelligence

30 years Business Intelligence: from Data Analytics to Big Data

Isabelle Linden
NADI institute, FOCUS research group
University of Namur, Namur, Belgium
Email: isabelle.linden@unamur.be

web-page: https://directory.unamur.be/staff/ilinden/

ABSTRACT

At the crossing of disciplines as Information Systems, Managment, Decision Support Systems, Data Mining, and Data Visualisation, Business Intelligence (BI) is understood in very different ways by the multiple concerned actors. This paper aims to offer to all of them an integrated view on multiple perspectives. To this end it first proposes a standard Business Intelligence approach. Then, it describes the main technical challenges addressed in the litterature with a particular focus on those risen by the emergence of Big Data.

Keywords: Business Intelligence, Big Data, Decision Support Systems

For a full version of this presentation please see this page at the end of the book.

DSS and BI: A bibliographic analysis of their links

Peter Keenan
University College Dublin
Dublin, Ireland
Email: Peter.Keenan@ucd.ie

web-page: https://people.ucd.ie/peter.keenan/

ABSTRACT

Decision Support Systems (DSS) and Business Intelligence (BI) are recognised as related disciplines and are often grouped together. However, different views exist as to the relationship between DSS and BI. Some authors characterise BI as a form of data-driven DSS, which suggests that BI is a subfield of DSS. Others suggest that BI has evolved beyond DSS or even replaced DSS, or simply ignore any connection with the DSS field,

This paper uses bibliographic techniques to examine the aggregate nature of the DSS and BI fields to understand the relationship between them. We investigated articles published in journals in the period 2000-2018 and the citations in those articles. This research identified that DSS is concerned with a wide range of applications, with significant clusters outside traditional areas in environmental and medical applications. The BI papers are much more closely associated with the traditional Information Systems (IS) field and generally cite papers from Computer Science or Business. The analysis showed that two journals *Expert Systems with Applications* and *Decision Support Systems* were important for both DSS and BI publications. DSS appeared in many specialist journals of no interest to the BI field, while a significant number of BI papers appear in BI journals disconnected from the DSS field. The research shows that BI is connected to the general Information Systems field and not only DSS, but that it is not as diverse as DSS.

Keywords: Decision Support Systems, Business Intelligence, Bibliographics

INTRODUCTION

Decision Support Systems (DSS) and Business Intelligence (BI) are recognised as related disciplines and are often grouped together, for instance in the ninth edition of the popular DSS textbook (Turban, Sharda, & Delen, 2011) and academic courses based on it. However, the exact nature of the connection between DSS and BI is not well established. Some authors characterise BI as a form of data-driven DSS (Power, 2008), which suggests that BI is a subfield of DSS. Others authors suggest that BI has evolved beyond (Watson, 2017) or even replaced DSS, while many authors simply ignore any connection with the DSS field, for instance, the widely cited BI paper (Chen, Chiang, & Storey, 2012) only references the term "Decision Support System" in relation to the name of the journal.

With the introduction of computerised bibliographic databases, bibliographic techniques have developed to allow the examination of aggregate trends in academic publication. One of the most comprehensive bibliographic databases is Web of Science (WOS) maintained by Clarivate Analytics (previously Thomson Reuters). This database records articles from 1898 to the present drawn from a wide range of disciplines and identifies the publications cited by those articles. WOS has close to 150 million records from 33,000 journals, with several billion cited references in the database (http://wokinfo.com). This paper uses bibliographic techniques to examine the aggregate nature of the DSS and BI fields in order to better understand the relationship between them. The disciplinary areas of application of DSS and BI are shown by the journals in which this research appears and the academic influences on these fields are shown by the journals cited by these papers. In addition, the keywords included with these articles indicate the themes of these papers.

In this research, we examined journal articles in WOS for the period 2000-2018 inclusive with the keyword "*Decision Support System*" or "*Business Intelligence*" in the title or topics and examined the papers cited by these papers. We did not investigate papers in conference proceedings and book chapters, which are less consistently represented in the WOS database than journals, making analysis more difficult. However, our analysis does identify citations to books or conference proceedings.

Table 1 Data extracted from the Web of Science.

Number	"Decision Support System"	"Business Intelligence"
Number. of Citing papers	8024	1329
Total number of cited papers	296724	58273
Number.of distinct cited papers	187526	39862
Number of distinct cited journals	2163	655

DECISION SUPPORT SYSTEMS

There were 6 times as many DSS papers as BI papers (Table 1) and consequently a larger number of papers and journals cited. In recent years DSS publications appear in a wide range of journals (Keenan, 2016) with some of the most highly ranked journals in the environmental or medical informatics domains and this is reflected in the journals with the largest number of DSS articles shown in Table 2.

Table 2 Main journals where DSS was published 2000-2018

Journals	**Discipline**
Expert Systems with Applications	ORMS
Decision Support Systems	ORMS
Environmental Modelling & Software	Environmental
Computers and Electronics in Agriculture	Environmental
European Journal of Operational Research	ORMS
Water Resources Management	Water Resources
Interfaces	Management
Journal of Environmental Management	Environmental
International Journal of Production Research	ORMS
Knowledge-Based Systems	Computer Science
Journal of Medical Systems	Medical Informatics
International Journal of Medical Informatics	Medical Informatics

In addition to examining the pattern of article publication, we can also look at the pattern of publication of papers cited by these DSS articles. The pattern of citations from DSS articles also reflects the wide range of disciplines addressed by DSS. We used the Vosviewer software (van Eck & Waltman, 2010, 2017) to visualise the co-citation network for the citations of DSS articles. Figure 1 shows the main journals cited by DSS papers and these cluster into four groups, a central cluster with long-established DSS journals including *Decision Support Systems* and *Expert Systems with Applications*, a cluster towards the top with OR/MS and operations management journals, a cluster on the left with environmental journals and a cluster with medical journals on the right.

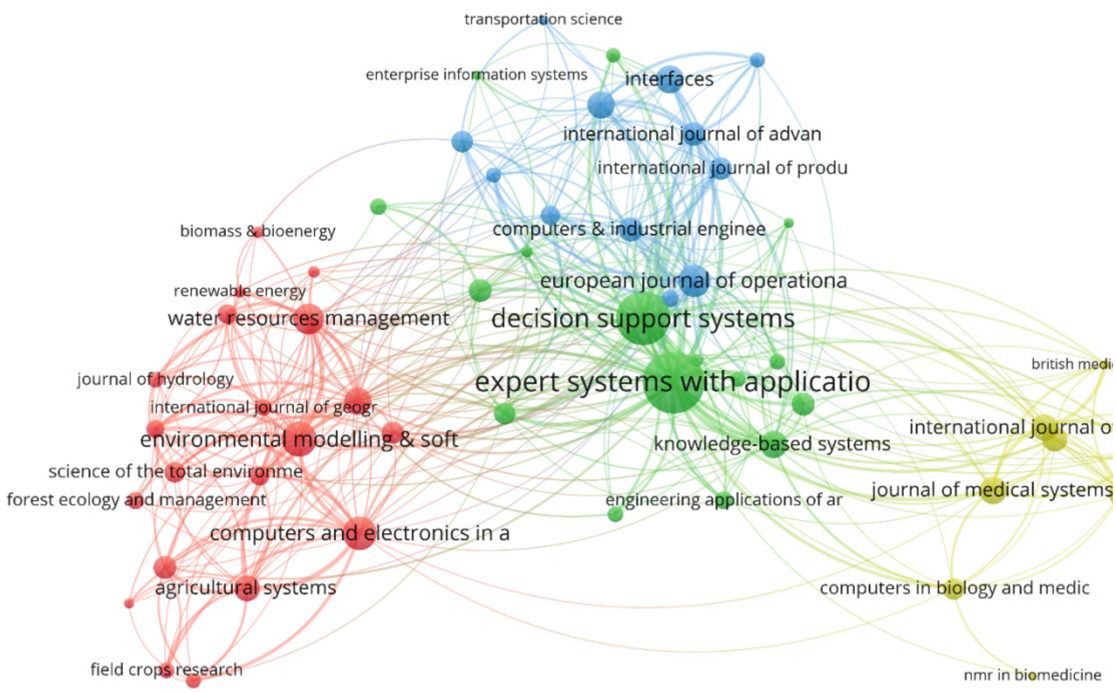


Figure 4: most important publications cited by BI papers visualised in Vosviewer

The BI domain has a smaller number of papers in the period 2000-2018. All of the journals with significant numbers of BI papers are in the traditional Information Systems (IS) domain (Table 3). Of these two journals, *Expert Systems with Applications* and *Decision Support Systems* are also important as an outlet for DSS publication. However, four of the top

Table 3 important journals for publication of BI papers.

BI Journal	**Importance in DSS**
Expert Systems with Applications	High
Decision Support Systems	High
International Journal of Information Management	Low
Journal of Intelligence Studies In Business	Low
Information Systems Management	Low
Journal of Computer Information Systems	Medium
Knowledge-Based Systems	Medium
Information Systems Frontiers	Medium
Industrial Management & Data Systems	Medium
International Journal of Data Warehousing and Mining	Low

ten BI journals have little or no relevance for DSS publication; *Journal of Intelligence* Studies in Business, *Information Systems Management*, *International Journal of Data Warehousing and Mining* and *Business & Information Systems Engineering.*

The WOS classifies journals in different ways, one classification is research areas. A particular journal may be classified into more than one research area, for instance both medical informatics and computer science. Figure 2 shows the proportions of DSS and BI papers published in journals in the most common research areas and this clearly shows the greater concentration of BI in the IS/Computer Science area and the diversity of DSS articles in fields such as Agriculture.

A similar picture emerges if we look at the papers cited by the BI articles, cited papers are also generally within the business or computer science domain. In Figure 3 the main journals cited can be grouped into BI related journals on the right, other IS journals on the left and other business journals towards the bottom of the figure. While there are a few references to medical informatics journals, these are still uncommon in BI papers despite their importance in DSS.

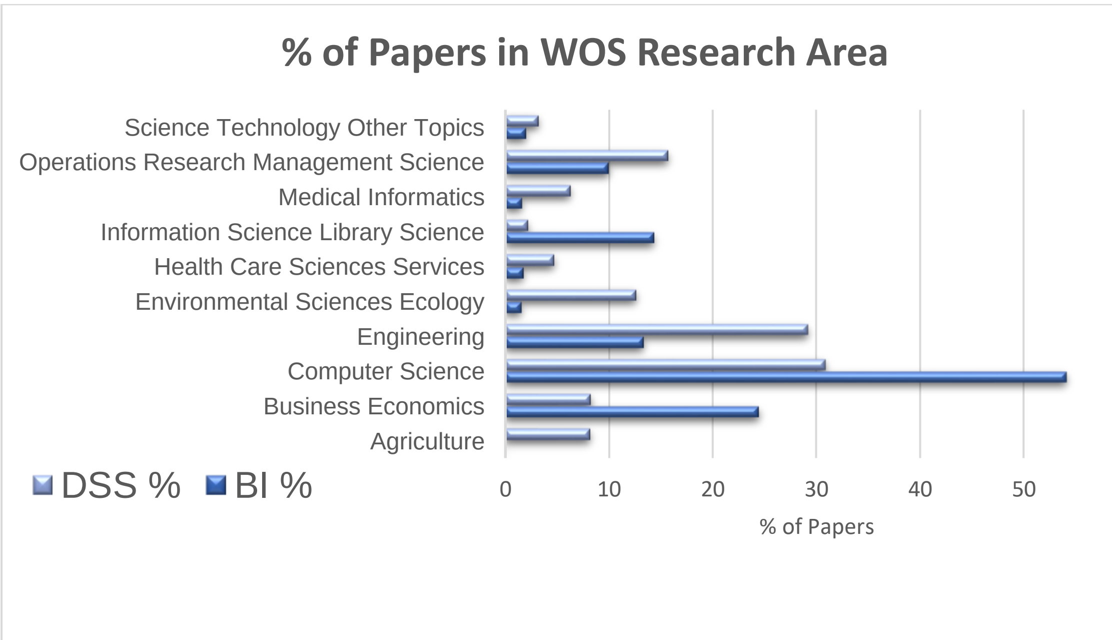


Figure 2: % of papers in main WOS Research Areas (note that categories overlap)

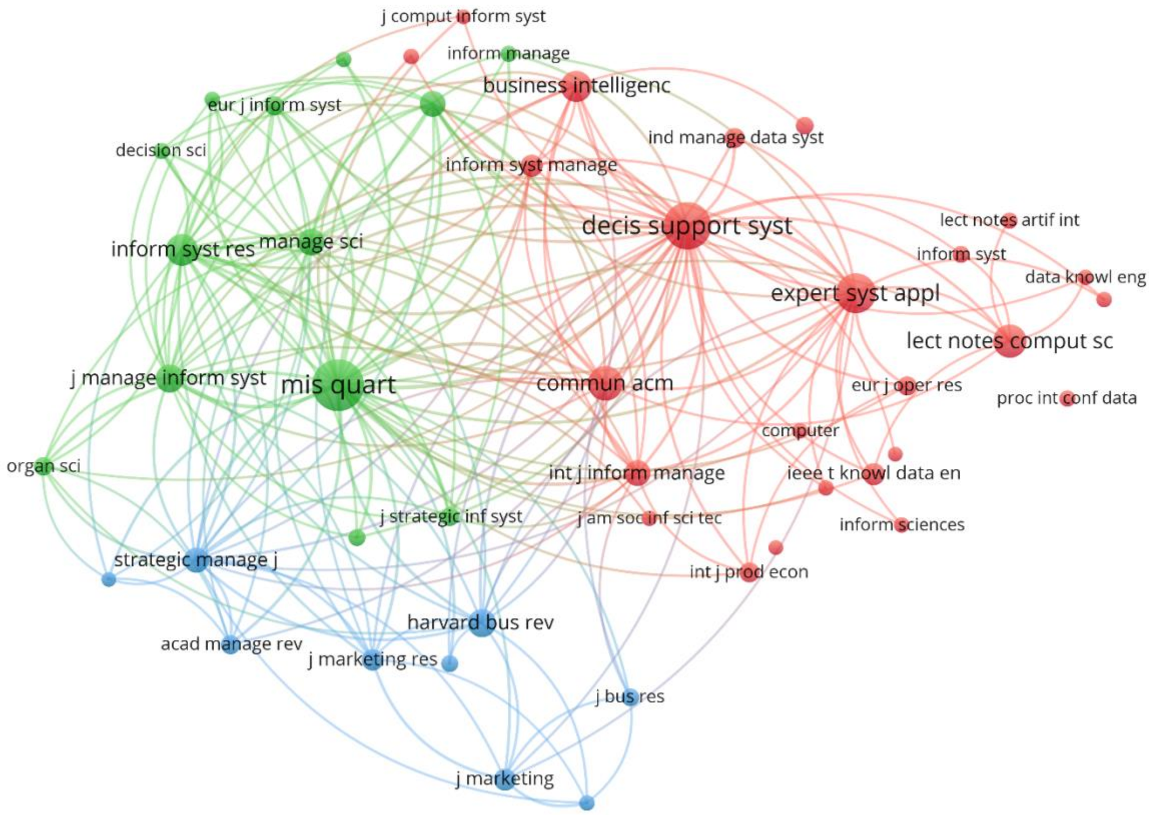


Figure 3 : most important publications cited by BI papers

In individual papers cited, the BI domain has some citations to older semimanual DSS papers and to some widely cited DSS review papers (Shim et al., 2002). There is also significant citation from both domains to some older IS papers, for instance (Davis, 1989). In terms of modelling approaches, Zadeh (1965) is often cited in both the BI and DSS domains, and there are a significant number of references to the TOPSIS method in both domains (Wang & Elhag, 2006), but Saaty's work is mainly found in the DSS domain only (Saaty, 1980). The design science approach (Hevner, March, Park, & Ram, 2004) is referenced significantly more commonly in the BI field than the DSS domain, reflecting commentary that this approach should be used more in DSS.

CONCLUSIONS

While this subject deserves further investigation beyond this short paper, we can nevertheless draw some conclusions from the data analysed. BI is a domain is strongly connected to the IS field generally, but not only to DSS. BI remains centred on business applications, whereas DSS is concerned with an increasingly wide range of applications, notably in the environmental and medical areas. However, model driven DSS continues to be relevant in specialist roles in traditional areas of application like business and engineering, BI has not "replaced" DSS. BI allows users to make sense of the increasing amounts of data in modern systems, using increasingly sophisticated analytics techniques, but this modelling is not typically manipulated by the user. DSS continues to support specialist applications where the decision maker often plays a direct role in building the solution through the use of the DSS. BI provides business value for a large proportion of business activity and so is important to business. Academic research concentrates on the more complex and novel problems and many of these are located outside the business domain. If the DSS field is too narrowly focussed on traditional application areas then it may appear to have become less important relative to BI, but this view reflects a limited perspective. However, it is probable that other disciplines who have adopted technology more recently may evolve in a similar way as IT use in business, only much more quickly. For instance, the concept of the "smart city" will involve large amounts of data being presented to city authorities and public alike and will synthesise BI techniques with environmental data (Pick, 2017). Business intelligence in the health domain is largely concentrated on financial and operations issues, but with increasing data coming from sensors and other devices patient data will be increasingly processing using BI techniques.

REFERENCES

[1] Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly, 36*(4), 1165-1188.

[2] Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly, 13*(3), 319-340. doi:10.2307/249008

[3] Hevner, A., March, S. T., Park, J., & Ram, S. (2004). Design science research in information systems. *MIS Quarterly, 28*(1), 75-105.

[4] Keenan, P. (2016). Changes in DSS disciplines in the Web of Science. *Journal of Decision Systems, 25*(Sup1), 542-549.

[5] Pick, J. B. (2017). Smart cities in the United States and worldwide: A rich arena for MIS studies. *Journal of Information Technology Case and Application Research, 19*(3), 133-144.

[6] Power, D. J. (2008). Understanding data-driven decision support systems. *Information Systems Management*, *25*(2), 149-154.

[7] Saaty, T. L. (1980). *The Analytic Hierarchy Process: planning, priority setting, resources allocation*. New York: McGraw.

[8] Shim, J. P., Warkentin, M., Courtney, J. F., Power, D. J., Sharda, R., & Carlsson, C. (2002). Past, present, and future of decision support technology. *Decision Support Systems, 33*(2), 111-126.

[9] Turban, E., Sharda, R., & Delen, D. (2011). *Decision support and business intelligence systems* (9th ed.). Boston: Prentice Hall.

[10] van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics, 84*(2), 523-538.

[11] van Eck, N. J., & Waltman, L. (2017). Citation-based clustering of publications using CitNetExplorer and VOSviewer. *Scientometrics, 111*(2), 1053-1070. doi:10.1007/s11192-017-2300-7

[12] Wang, Y.-M., & Elhag, T. M. S. (2006). Fuzzy TOPSIS method based on alpha level sets with an application to bridge risk assessment. *Expert Systems with Applications, 31*(2), 309-319. doi:https://doi.org/10.1016/j.eswa.2005.09.040

[13] Watson, H. J. (2017). Preparing for the Cognitive Generation of Decision Support. *MIS Quarterly Executive*.

[14] Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*(3), 338-353.

A conceptual business process entity with lifecycle and compliance alignment

George Tsakalidis, Kostas Vergidis, Pavlos Delias, Maro Vlachopoulou
Department of Applied Informatics,
University of Macedonia, Thessaloniki, Greece.
{giorgos.tsakalidis,kvergidis,mavla}@uom.edu.gr

Pavlos Delias
Department of Accounting and Finance, Eastern Macedonia and Thrace Institute of Technology, Kavala, Greece
pdelias@teiemt.gr

ABSTRACT

This paper proposes a conceptual model that incorporates: (a) a business process entity, (b) a business process lifecycle aligned with the proposed entity, and (c) a compliance framework that focuses on the degree of process compliance to imposed regulatory standards. The pursuing objective is to systematize business processes through a conceptual entity applicable to Business Process Management (BPM) practices and compliance-checking. The applied methodology involves the review, interpretation and comparison of business process definitions, structural elements and their interrelations, acknowledged BPM lifecycles and compliance rules. The initial findings lead to the proposal of a contextual business process structure that sets the boundaries of business process as a clearly defined entity. The business process entity encompasses continuous modification of its design based on the feedback it generates. Additionally, a comparative analysis of prominent BPM lifecycles resulted in a proposed 'business process lifecycle' that allows for a better alignment of the included cycle steps. The proposed business process entity is also related with process compliance practices to produce compliance-aware business processes. The introduced conceptual model can assist professionals in apprehending core business process features, focusing on process flexibility and redesign. It can also serve as a preliminary prototype for checking the degree of compliance between a business process and the applicable compliance rules.

Keywords: Business Process, Business Process Management, BPM lifecycle, Business Process Lifecycle, Business Process Compliance.

Find and cite the full work as:
Tsakalidis, G., Vergidis, K., Delias, P., Vlachopoulou, M., 2019. A conceptual business process entity with lifecycle and compliance alignment, in: Freitas, P.S.A., Dargam, F., Moreno, J.M. (Eds.), Decision Support Systems IX: DSS Developments & Future Trends, Lecture Notes in Business Information Processing. Springer International Publishing, Cham, pp. 65-77.

Two Problem Situations for Process Innovation Based on Operations Sophistication

Pavlos Delias
Eastern Macedonia and Thrace Institute of Technology
Agios Loukas, Kavala, PC 65100, Greece
pdelias@teiemt.gr

Daniela Grigori
Univ. Paris-Dauphine, PSL Research University, CNRS UMR[7243], LAMSADE, 75016
Place du Maréchal de Lattre de Tassigny, 75016 Paris, France
Daniela.grigori@dauphine.fr

ABSTRACT

Process Innovation is assumed to require a more intrinsic rethinking of business processes, which is typically a creative process. Nevertheless, in this creative, prolific process there can be artifacts derived from rational practices that are capable to provide insightful recommendations. In this work, we claim that an event log, a file that registers the execution of the relevant business processes, can be the source of such an artifact.

We describe the fundamental elements of two problem formulations, namely the set of alternatives, the set of potential actions that the decision maker may undertake, the set of points of view (dimensions) from which the potential actions are observed, analyzed, evaluated, compared, etc, and the problem statement (what is expected to be done with the alternatives) for two cases. The one refers to the case when multiple organizations perform the same business process (a case commonly taking place in public organizations, or in multinational companies), and the second one refers to when multiple resources are involved by performing similar roles in a single business process.

In both cases, the event log provides the input to create a bipartite network whose structure reveals the operations sophistication, which in turn exposes the process innovation potentials. Although we briefly discuss how this is possible, in this work we emphasize the problem formulation rather than the evaluation technique that exploits the event log.

Keywords: Process Innovation, Problem Formulation, Process Mining, Evaluation Models

INTRODUCTION

The seminal work of [1] identifies process innovation within the core element of *Methods* as the stage that includes all methods which facilitate the development of improved business processes, as well as within the core element of *Information Technology* as the (semi-)automated support for the generation of improved business processes. There are of course different types of process innovation and a plethora of approaches towards it, however, arguably, among the factors that can stimulate and promote it, we find the reduction of innovation latency [2] (i.e., to timely anticipate what improves the process), and the mitigation of the organizational resistance towards a process change endeavor [3].

In this work, we propose a way to address both these issues under two different problem situations. We claim that this can be achieved by looking and taking action according to a metric that we call *Operations Sophistication*. This metric reflects how much diversified are the capabilities of an organization, and how difficult it is to deploy and apply those capabilities to the organizational operations. The core assumption is that a more versatile organization is expected to be able to demonstrate higher sophistication, manifest more complex sets of capabilities, and hence achieve better performance.

More specifically, we assume a tripartite network. The one part consists of the capabilities, the second of the organizations, and the third of the behaviors/ patterns they exhibit over their operations within a business process. Organizations are connected to their available capabilities, and patterns are connected to their required capabilities. Then, if a pattern is observed, we can assume that all the required capabilities are in place, therefore, we can reflect the tripartite network with a bipartite one, connecting organizations to patterns.

In previous work [4], we have shown that by analyzing the structure of such a bipartite network (an organizations-patterns network) and by iteratively considering the properties of the neighboring nodes of each part, it is possible to derive several metrics that reveal the potentials of the patterns to contribute to higher sophistication for the organizations. However, even after deriving several relevant metrics, it is not clear how a problem situation can be shaped and how a corresponding decision model can be formulated. How organizations should proceed? What is the shape of a solution? Why such a solution should be preferred against another? In this context, this work contributes by defining two problem situations that eventually allow the identification of the underlying structure of the relevant decision problem, and as such, they allow the re-use of procedures and models [5] forging a consistent process innovation technique.

In the next section we try to give an overview of the concepts related to the Operations Sophistication notion, while in section 3, we present the two problem situations along with all the required definitions. We acknowledge that the presentation is not very descriptive, but this is due to the nature of this work, and we expect to elaborate on these aspects in future work.

BACKGROUND

In this section we briefly describe the related concepts and terms. Since only the basic justifications and information are provided, the interested reader is redirected to [4], [6] for a more detailed description. First, we assume that there is an available event log [7, p. 128] which contains the execution logs of the business process for which the innovation project is initiated. In [4], we describe analytically how this event log can produce a bipartite network. This bipartite network can take two shapes, as we will explain in the next section. The one set of nodes will always represent patterns (process behaviours like batch processing, involving multiple resources, balancing workload, etc. that are routinely followed during the process execution). The second set will reflect either the resources that are performing the relevant process tasks, or it will comprise organizations that are performing the same process. Without loss of generality, from here on we will describe the concepts based on the second option.

We define *versatility* as the number of the different patterns that an organization demonstrates. In addition, we define the *pervasiveness* of a pattern as the number of organizations that exhibit it. We expect that patterns that require complex combinations of capabilities to be less pervasive, or in other words, the pervasiveness of a pattern signals the amount of capabilities that are required for its application. Therefore, we claim that the performance of an organization resides in the diversification of the patterns it can demonstrate. Sophistication is expressed in the assembly of the exhibited patterns and reflects the mechanisms that are needed to exploit different sets of capabilities.

By using versatility to correct pervasiveness and vice versa, through a set of recursive equations (see [6]) we can capture a metric that reflects the degree of sophistication that each organization exhibits during the process execution (we shall call this the *Operations Sophistication Index-OSI*) and a similar metric that reflect how much sophisticated every pattern is (let us call this *Patterns Sophistication Index – PSI*).

Using the structure of the bipartite network, as well as the above basic metric, we can ramify the outputs and construct a set of additional insightful metrics. By calculating the probability of co-occurrence in the same organization between every two patterns, we can calculate a kind of similarity (*proximity)* over them. Then, exploiting this proximity metric among patterns, we can calculate the distance of every organization to every pattern as the normalized sum of the proximities of all the patterns that are proximate to the relevant pattern but that are not yet exhibited by the relevant organization.

Accounting not only for the distance between an organization and patterns but for the level of patterns' sophistication as well, we could calculate the sophistication prospects of the organization. This kind of weighting, results in a metric which we call *Opportunity Value* for an organization and it is very important, because when we create differences between the opportunity value of an organization and its alter ego that does not apply a distinct pattern, we can calculate the *Opportunity Gain* that every organization has from the adoption of that distinct pattern.

TWO PROBLEM SITUATIONS

In this section, and following [5], we will take the view that a decision aiding process is a process "*in which different agents endowed with cognitive capabilities have to share some information and knowledge in order to establish some shared representation of the process object*". During the first steps of the process, these pieces of information of knowledge take the shape of two major deliverables, namely the problem situation and the problem formulation. The former boils down to a representation that will ultimately aid the client to better arrange herself regarding the decision procedure for which she asked the analyst's recommendation. The latter (problem formulation), is actually a task of translating the client's interest into a format that decision support techniques and methods can address. This is reached by using a formal decision support language, however since this will inevitably lead to a reduced reality, we ought to point out the following pitfalls: A problem formulation is not neutral to the final recommendation (solution), indeed a different formulation is very likely to lead to a different recommendation. The analyst's defense of this is that following a problem formulation, the client will eventually be able to anticipate the possible conclusions and check whether these are compatible with her expectations. It is quite clear that the analyst shall not continue the decision aiding process, unless the problem formulation is validated by the client.

In this work, and in accordance with [5], [8], we define a problem situation $\mathcal{P}$ to be a triplet $\mathcal{P} = \langle \mathcal{A}, \mathcal{O}, \mathcal{S} \rangle$ where $\mathcal{A}$ are the actors involved in the process, $\mathcal{O}$ are the objects (problems, interests, opportunities) introduced by each actor (for instance a manager of an organization may be concerned with the organizational change that an innovation will bring, while shareholders may be concerned with the expected profits it will bring), and $\mathcal{S}$ are the resources (monetary or not) committed by each actor to each object of her concern. Another triplet representation is employed for the problem formulation. In particular, we define a problem formulation ΓΓ to be a triplet $\Gamma = \langle \mathbb{A}, \mathbb{V}, \Pi \rangle$, where $\mathbb{A}$ is the set of alternatives, the set of potential actions that the client may undertake, $\mathbb{V}$ is a set of points of view (dimensions) from which the potential actions are observed, analyzed, evaluated, compared, etc, and Π is the problem statement (what is expected to be done with the elements of $\mathbb{A}$ - some common problem statements are choice, ranking, rejection, etc.).

Formulation at the level of organizations

This is the case where multiple organizations are performing the same business process (in parallel, not collaboratively), and they are interested in improving their relevant performance. This scenario is typical in the case of public organizations, or in branches of multinational companies. Following the definitions of the previous section, we postulate that:

- $\mathcal{A}_{org}$ is the set of actors that will get affected by the consequences of the decision, as well as the actors that are influencing the decision. This set comprises *i*) representatives from each organization: the CEO; the process owner (the person who is responsible for the efficient and effective operation of the relevant business process); the process participants (the persons that are performing the process) and *ii*) the process analyst; the process methodologist; a higher authority that is the actor who initiate the process innovation project.
- $\mathcal{O}_{org}$ comprise a performance metric which is the target of improvement, the process behaviors that are routinely followed by the process participants during the process execution, and the organizational resistance to change.
- $\mathcal{S}_{org}$ consists besides of the implicit elements (the labour, the knowledge, etc.), of an event log (a flat file that registers the traces of the process execution) and of a set of *patterns*, i.e., process behaviors (e.g., batch processing, workload balancing) which organizations demonstrate during the process execution.
- $\mathbb{A}_{org}$ is a set of a patterns that can be adopted in order to improve the performance of organizations. We should emphasize two important issues: *i*) we assume that the process innovation project consists of the adoption of a subset of those patterns for the process execution; *ii*) patterns are meant as the elements of the final solution as a reflection of the capabilities that are required to perform them.
- $\mathbb{V}_{org}$ the evaluation dimensions include the potentials of improvement that each pattern brings, the feasibility of its realization in terms of the organizational resistance that it will meet, and the suitability of the pattern for the particular organization.
- Π_{org} is the problem statement which we suggest being non-purposeful, i.e., to describe the patterns under the points of view established in $\mathbb{V}_{org}$.

Formulation at the level of human resources

In this case, we twist the main idea, and modify the objects of the bipartite network. Actually, we modify just the objects of the one set of nodes, the one with the organizations. In this, version, the bipartite network will therefore contain human resources (instead of organizations) and patterns. More specifically, consider the case when multiple human resources in an organization are assuming similar work roles. Being participants in the same business process, the human resources are interested in improving their personal competencies with respect to their performance about the process. The formulation elements for this case are:

- $\mathcal{A}_{hr}$ is the set of the process participants, the human resources manager (or any other manager that performs some kind of performance appraisal), and the process owner, in the sense that he/she provides the process performance metrics and evaluations.
- $\mathcal{O}_{hr}$ comprise a performance metric which is the target of improvement, the atomic behaviors that are routinely followed by the process participants during the process execution, and the motivation (including incentives) of each participant to improve her own personal performance.
- $\mathcal{S}_{hr}$ consists besides of the implicit elements (the labour, the knowledge, etc.), of an event log (a flat file that registers the traces of the process execution) and of a set of

patterns, i.e., atomic behaviors (e.g., extreme handover, intense collaboration) which the process participants demonstrate during the process execution.

- $\mathbb{A}_{hr}$ is a set of a patterns that can be adopted in order to improve the performance of every individual participant. We should emphasize two important issues: *i*) we assume that the process innovation project consists of the convincing participants to adopt a subset of those patterns so that eventually the process execution is improved; *ii*) patterns are meant as the elements of the final solution as a reflection of the capabilities that are required to perform them.
- $\mathbb{V}_{hr}$ the evaluation dimensions include the potentials of improvement that each pattern brings, the feasibility of its realization in terms of the possession of the required personal capabilities, and the suitability of the pattern for the particular resource.
- Π_{hr} is the problem statement is similar with Π_{org} since we suggest again to be non-purposeful, i.e., to describe the patterns under the points of view established in $\mathbb{V}_{hr}$.

CONCLUSIONS

In this work we described how a decision model can be formulated for two different situations. In both cases, the triggering object is an event log which registers the execution of a business process, and the expected outcomes have the shape of recommendations for pertinent improvements. In one case, the situation is centered at the level of organizations that perform the same process and strive for a growth upbeat, and in the second case, the situation is centered at the level of human resources that strive for amelioration and self-development.
There are of course many additional issues that were not discussed in this work, especially about the role that the event log plays, as well as about the validation options of the operations sophistication approach. However, this work consists a part of an ongoing work, that is partly shaped by the problem situations defined here, and as such its capacity becomes essential.

REFERENCES

[1] M. Rosemann and J. vom Brocke, “The Six Core Elements of Business Process Management,” in *Handbook on Business Process Management 1: Introduction, Methods, and Information Systems*, J. vom Brocke and M. Rosemann, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 105–122.

[2] M. Rosemann, “Proposals for Future BPM Research Directions,” in *Asia Pacific Business Process Management*, vol. 181, W. van der Aalst, J. Mylopoulos, M. Rosemann, M. J. Shaw, C. Szyperski, C. Ouyang, and J.-Y. Jung, Eds. Cham: Springer International Publishing, 2014, pp. 1–15.

[3] W. M. P. van der Aalst, M. La Rosa, and F. M. Santoro, “Business Process Management: Don’t Forget to Improve the Process!,” *Business & Information Systems Engineering*, vol. 58, no. 1, pp. 1–6, Feb. 2016.

[4] P. Delias, M. Acheli, and D. Grigori, “Applying the Method of Reflections through an Event Log for Evidence-based Process Innovation,” presented at the 1st International Conference on Process Mining, Aachen, Germany, 2019.

[5] D. Bouyssou, T. Marchant, M. Pirlot, A. Tsoukiás, and P. Vincke, *Evaluation and decision models with multiple criteria: stepping stones for the analyst*. New York: Springer, 2006.

[6] C. A. Hidalgo and R. Hausmann, “The building blocks of economic complexity,” *Proceedings of the National Academy of Sciences*, vol. 106, no. 26, pp. 10570–10575, Jun. 2009.

[7] W. van der Aalst, *Process Mining: Data Science in Action*, 2e ed. Springer-Verlag Berlin Heidelberg, 2016.

[8] I. Stamelos and A. Tsoukiàs, "Software evaluation problem situations," *European Journal of Operational Research*, vol. 145, no. 2, pp. 273–286, Mar. 2003.

[9] P. Delias and N. Spanoudakis, "Simulating multi-agent system designs using business process modeling", in: Proceedings of the 8th European Workshop on Multi-Agent Systems (EUMAS 2010), Paris, France, December 16–17 2010.

Special Talks

DSS in Complex Environments

Ana Isabel Barros
Principal Scientist at TNO, Netherlands Organisation for Applied Scientific Research
Fellow at the University of Amsterdam, Institute for Advanced Study (IAS)
Senior Research Fellow at the Netherlands Defence Academy

The on-going globalisation and increasing accessibility of information, due to advances in information and communication technology, yield an increase in societal connectivity and dynamics as well as a better perceived shared awareness and quicker access to information and knowledge. This, combined with other mega-trends such as the rapid rate of technological development (e.g. Artificial Intelligence), climate changes (leading to extreme weather events), demographic changes (e.g. large-scale involuntary migration) and increasing economic inequality yield a complex and very interconnected society.

In such a complex environment decision making and its support faces several challenges. Firstly, how to integrate the constant feed of huge amounts of data into the decision making in a meaningful and understandable manner? Secondly, how to cope with the pressure of making decisions quickly due to the increasing pace of life and the instant visibility of decisions? Finally, how to assess the effects of decisions (not only first order but also higher order) for different time horizons and for different partners and stakeholders? This talk will explore how these decision support challenges could be tackled in practice.

Natural Language Processing in the SSIX project

Brian Davis
Lecturer (ATB) in Computer Science at Maynooth University, Ireland

Social Sentiment Indices powered by X-Scores (SSIX) is a recent successfully completed Horizon 2020 innovation action; its mission was to provide European SMEs with a collection of easy to interpret tools to analyse and understand social media users attitudes for any given subject; these sentiment characteristics can be exploited to help SMEs to operate more efficiently resulting in increased revenues. Social media data represents a collective barometer of thoughts and ideas touching every facet of society. SSIX searches and indexes conversations taking place on social network services, such as Twitter, StockTwits, and Facebook including the most reliable and authoritative Newswires, online newspapers, trade publications and blogs. SSIX classifies and scores content using a framework of qualitative and quantitative parameters called X-Scores, regardless of language, locale or data architecture. The X-Scores framework interprets economically significant sentiment signals in social media conversations producing sentiment metrics, such as momentum, awareness scope, topic density and historical comparison. These metrics can create commercially viable social sentiment indices, which can be tailored to any domain of interest. By enabling European SMEs to analyse and leverage social sentiment in their discipline, SSIX facilitates the creation of innovative products and services by enhancing the investment decision making process, thus increasing revenue while also minimising risk. This talk will focus on the applied research output of the project with respect multilingual aspect orientated sentiment analysis from financial microblogs. I will also discuss the challenges encountered in the SSIX project with respect to processing social media content, the various levels of linguistic analysis required to overcome these and solutions and resources provided by the project.

Making Business Meaning from Bits & Bites - 3rdPLACE & External Data Intelligence (E.D.I.)

Pierluigi Vacca
3rd PLACE – Partner & Chief Product & Service Officer, Italy.

The most effective way for 3rdPLACE to exploit the SSIX experience from a commercial perspective was to shift from a product-based outcome to a service-based outcome named E.D.I.

What is EDI?

Mission: to turn huge volume of unstructured data into business meaning for mid-big company. The output is represented by studies, papers, researches, etc.

How? leveraging both technology (as the SSIX Framework) and methodology to provide an alternative set of digital insight

Why? To enrich the informational base to support a better decision making on different goals as: launch of new product /service, open new international markets, consolidate the brand in the domestic market, etc.

From a technical point of view, the SSIX experience concurred to give 3rdPLACE both knowledge and tech skills in big data collection in NLP techniques as well in building internal algorithms and solutions.

Business case

In 2017, we had the chance to apply this new service to a real business case. Infact we were approached by the italian country manager of Rothschild & CiE Gestion (one of the world's largest independent financial advisory groups). In particular he needed an analysis aimed to collect public data from the web to measure and interpret users' interest towards wealth management. The main goal was to leverage a data-driven approach to stimulate a digital transformation in financial advisory

In particular, the most relevant questions to be answered were:

- Do users talk, ask questions, search for assistance, share experiences online about wealth management?
- If and How much do these online conversations grow year by year?
- Where is this interest mostly concentrated? Social networks? Forums? Blogs?
- What are the most discussed topics and subtopics related to wealth management? (transparency, robs-advisors?)
- What is the user's sentiment related to those topics and subtopics?

Leveraging the SSIX platform, 3rdPLACE could:

- Collect nearly one year of data coming from different digital sources (Twitter, finance blogs, RSS feeds, business information news, etc), (some data were bought from 3rd party companies)
- Run data sampling and NLP algorithms to reduce the noise and to evaluate only relevant contents or source partnerships for more complex tasks,
- Categorize the data samples into topics and subtopics,
- Measured the user's sentiment on each topic/subtopic.
- Visualises insights through third parties application.

The final outcome ended up to be a big success. Basically this paper (print some copies) demonstrated that exist an online demand for wealth management (users ask questions, need information and support and share experience). This demand represent an opportunity to all the consultants that are open to embrace and adopt the digital medium. From this point of view we were invited to a roadshow up&down Italy to present this paper to the financial advisors.

Conclusions

3rdPLACE's mission has always been to "turn data into business" but, before SSIX project, we have always collected and analyzed only our clients internal data. The SSIX Consortium gave us the opportunity to broaden our capacity to work with data, getting precious know how to manage also huge volume of public data at scale. This know how represent today an invaluable asset for 3rdPLACE to support its client not only to derive value from their proprietary data but also to turn public data into actionable stories to optimize their business.

Sustainable Developments & Approaches

Value-chain wide food waste management: a systematic literature review and recommendations for future work

Guoqing Zhao, Shaofeng Liu, Huilan Chen, Carmen Lopez
University of Plymouth, Drake Circus, PL4 8AA, Plymouth, UK
{guoqing.zhao;shaofeng.liu;huilan.chen;carmen.lopez}@plymouth.ac.uk

Jorge Hernandez
University of Liverpool, Liverpool, L69 3BX, UK,
2 J.E.Hernandez@liverpool.ac.uk

Cécile Guyon
Bretagne Développement Innovation, 1bis Rue de Fougères, 35510,Cesson-Sévigné, France,
c.guyon@bdi.fr

Rina Iannacone
ALSIA, Viale Carlo Levi, 75100 Matera MT, Italy,
rina.iannacone@gmail.com

Nicola Calabrese
CNR-ISPA, Via Amendola, 122/O – 70126 Bari BA – Puglia, Italy
nicola.calabrese@ispa.cnr.it

Hervé Panetto
University of Lorraine, 34 cours Léopold, CS 25233, F-54052 Nancy cedex, France,
herve.panetto@univ-lorraine.fr

Janusz Kacprzyk
IBSPAN, Newelska 6, 01-447 Warsaw, Poland,
janusz.kacprzyk@ibspan.waw.pl

MME Alemany
Universitat Politècnica de València, Camino de Vera, s/n, 46022, Valencia, Spain,
marvea@omp.upv.es

ABSTRACT

The agriculture value chain, from farm to fork, has received enormous attention because of its key role in achieving United Nations Global Challenges Goals. Food waste occurs in many different forms and at all stages of the food value chain, it has become a worldwide issue that requires urgent actions. However, the management of food waste has been traditionally segmented and in an isolated manner. This paper reviews existing work that has been done on food waste management in literature by taking a holistic approach, in order to identify the causes of food waste, food waste prevention strategies, and elicit recommendations for future work. A five step systematic literature review has been adopted for a thorough examination of the existing research on the topic and new insights have been obtained. The findings suggest that the main sources of food waste include food overproduction and surplus, food waste caused by processing, logistical inconsistencies, and households. Main food waste prevention strategies have been revealed in this paper include policy solutions, packaging solutions, date-labelling solutions, logistics solutions, changing consumers' behaviours, and reuse and redistribution solutions. Future research directions such as using value chain models to reduce food waste and forecasting food waste have been identified in this paper. This study makes a contribution to the extant literature in the field of food waste management by discovering main causes of food waste in the value chain and eliciting prevention strategies that can be used to reduce/eliminate relevant food waste.

Keywords: Systematic literature review, food waste generation, cause of waste, food waste prevention strategies

Find and cite the full work as:
Zhao, G., Liu, S., Chen, H., Lopez, C., Butel, L., Hernandez, J., Guyon, C., Iannacone, R., Calabrese, N., Panetto, H., Kacprzyk, J., Alemany M., 2019. Value-chain wide food waste management: a systematic literature review and recommendations for future work, in: Freitas, P.S.A., Dargam, F., Moreno, J.M. (Eds.), Decision Support Systems IX: DSS Developments & Future Trends, Lecture Notes in Business Information Processing. Springer International Publishing, Cham, pp. 39-51.

Assessing multidimensional sustainability of European countries with a novel, two-stage DEA

Georgios Tsaples, Jason Papathanasiou, Andreas C. Georgiou
Department of Business Administration, University of Macedonia, Egnatia Str. 156 54636, Thessaloniki Greece
gtsaples@gmail.com, jasonp@uom.edu.gr, jasonp@uom.edu.gr

Nikolaos Samaras
Department of Applied Informatics, University of Macedonia, Egnatia Str. 156 54636, Thessaloniki Greece
samaras@uom.edu.gr

ABSTRACT

The issues of sustainability and sustainable development are considered important aspects of governmental policy making. Currently, many methods are used to assess the sustainability of various regions and countries each accompanied by advantages and disadvantages. The objective of the current paper is to propose a new methodological framework for the assessment of sustainability that attempts to mitigate some of the limitations of the methods that are used. The proposed method is based on two-stage Data Envelopment analysis. In the first stage, raw data are transformed to sub-indicators using the multiplicative version of the VRS DEA model. The sub-indicators are used in the second-stage, in a typical Benefit-of-the-Doubt model, to calculate their optimal weights, which are used in the construction of a geometric composite indicator of sustainability. The method is tested to calculate a sustainability index for 28 European countries. The results show that eastern European and Scandinavian countries appear to be more sustainable than western, developed countries.

Keywords: Sustainability, Data Envelopment Analysis, Benefit-of-the-Doubt, composite indicator.

Find and cite the full work as:
Tsaples, G., Papathanasiou, J., Georgiou, A.C., Samaras, N., 2019. Assessing multidimensional sustainability of European countries with a novel, two-stage DEA, in: Freitas, P.S.A., Dargam, F., Moreno, J.M. (Eds.), Decision Support Systems IX: DSS Developments & Future Trends, Lecture Notes in Business Information Processing. Springer International Publishing, Cham, pp. 103-115.

UML-based Modelling of Information Systems for Serious Game Tracebility: Application to Decision Support for Avalanche Risk

Laure Vidaud
Irstea, UR ETGR, Université Grenoble Alpes
2, rue de la papeterie, 38402 Saint-Martin d'Hères, France
laure.vidaud@irstea.fr

François Pinet
Université Clermont Auvergne, Irstea, UR TSCF, Centre de Clermont-Ferrand
9 avenue Blaise Pascal CS 20085, F-63178 Aubière, France
francois.pinet@irstea.fr
www.irstea.fr/francois-pinet

Jean-Marc Tacnet
Irstea, UR ETGR, Université Grenoble Alpes
2, rue de la papeterie, 38402 Saint-Martin d'Hères, France
jean-marc.tacnet@irstea.fr
www.irstea.fr/tacnet

ABSTRACT

In this paper, we propose a design method based on the Unified Modeling Language (UML) to formalize information systems dedicated to trace decision making reasoning in serious games. The purpose is to propose a uniform way to represent several serious games. The main goal of our approach is to facilitate the conceptual modelling of the traceability of the stages of the game, the information provided to players and the choices made by them. We also describe a serious game application developed for decision making related to avalanche risk and apply our UML-based method on this use case.

Keywords: Serious Game, UML, Decision Support, Avalanche Risk

INTRODUCTION

- In this paper, we propose a design method based on UML (Unified Modelling Language) profiles to model information systems dedicated to trace decision making in a serious game. We also present a serious game application we developed dedicated to decision making related to avalanche risk. A first use of our UML approach is presented for this case study. In this application, the goal of our method is to facilitate the conceptual modelling of the traceability of the stages of the game, the information provided to players and the choices made by them. The result presented in this paper is a data model that can be used to store the traceability of the game sessions in a database. Then, the study of these data will allow the analysis of players' behaviours throughout the decision process. Our UML profile can be re-used for other serious game applications.

The paper is organized as follows. The second section describes the serious game we developed. The third section presents a brief bibliographic review on serious game design and the application of our serious game UML-based method on our game. A conclusion sums up our contribution and presents some perspectives.

A SERIOUS GAME FOR DECISION-MAKING RELATED TO AVALANCHE RISK

In this game, a team of players have to make a collective decision in a context of avalanche risks preceding (or not) a period of crisis management. The serious game we developed aims to put the players of the team in situation of decision. This decision is related to close (or not) a road exposed to one or several avalanche corridors. This game allows people to play different roles: a local elected representative, a road manager or an avalanche expert. There is no decision procedure planned in advance when the game is starting. The team of players have to construct its own reasoning step by step with each new piece of data arriving in the game. The research challenge is to observe and analyse the reasoning that led to the decision, in a context of more or less imperfect data.

We have proposed a UML-based model of the whole process leading to a decision. Our future goal will be to use this conceptual modelling to create a database to store every event that has occurred during gaming sessions. These data will be used to analyse the decision-making process.

The main game mechanics are easy to understand. In a first phase, the team learns the real-life facts based scenario of the game. Then, the game is composed of several rounds. These turns are all similar. In a final phase, more information on the scenario is provided to the players, including the real-life outcome. During the game, some of the information is given randomly, by card draws. Here are the details of the game.

1. **Game start:** different types of information are provided to all the players in order to understand the context: location of the site, weather situation, avalanche risk level, social context.
2. **Game turns**: each game turn is as follows:
    - The team can ask for a limited number of information, only one per turn. The team can choose the information type. Figure 1 shows the different categories of information provided.
    - The team obtains new information in the chosen category. The game system provides randomly different information quality level. The information provided to players is more or less imperfect (precise, vague, etc.). This new information is evaluated individually by each player and collegially by the team.
    - The team makes the decision to continue playing (obtain new information to refine

the situation analysis) or to make a decision on the road, which ends the game.

3. **Game conclusion**: The team has to prepare a justification of its choices to several types of stakeholder (prefect, insurance companies, etc.). At the end, a synthesis of the real facts that inspired the game scenario is provided and a pedagogical debate is initiated with the players.

At any time during the game turn phases, events can occur and the game context can change to improve or to deteriorate the social context of the game or the meteorological conditions of the phenomenon. These events, provided by the animator of the game, force the team to re-evaluate the situation. Three things can cause the end of a game: the information that the avalanche has fallen (i.e., it is too late to make a decision), a decision is made by the team, the playing time is up.

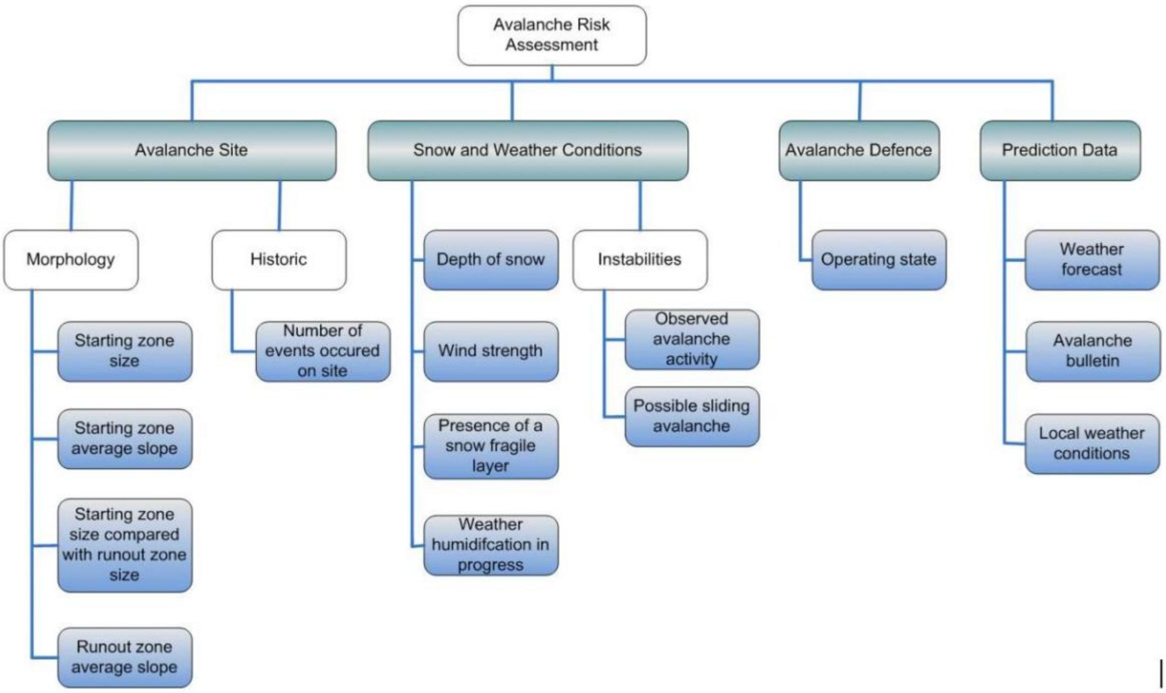


Figure 1. Possible information provided to players during the serious game

UML-BASED DESIGN FOR SERIOUS GAME TRACEABILITY

Serious game design methods have interested different researchers for some fifteen years [1]. Serious games have received renewed interest in recent years and the methods for designing them have been the subject of specific research. The main serious game design methods, based on model approaches, were briefly presented in [2], [3] and [4]. Design frameworks have also been proposed in [5], [6], [7]. More recently, object-oriented model have been proposed in [8], [9]. In the context of ontologies, a formal representation has been introduced in [4]. Design patterns dedicated to serious games have also been proposed in [10], [11], [12]. No integrated method specifically dedicated to the development of information system for the traceability of serious games has been proposed. We made a synthesis of these proposals in order to propose a UML profile for serious game traceability design. In this paper, we illustrate the use of our UML profile on the conceptual model of the serious game presented in the previous section, but our profile can be used for the modelling of traceability of other serious games.

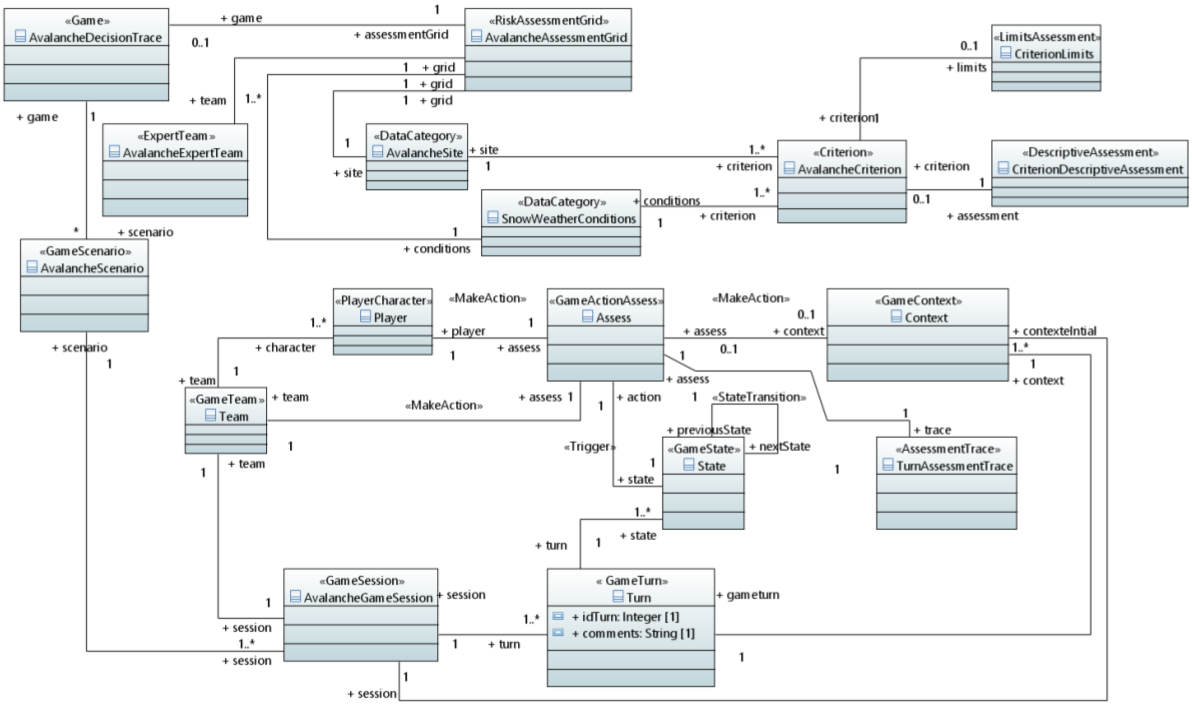


Figure 2. Example of use of our UML profile for serious game traceability modelling (part A)

Our UML profile offers a set of stereotypes, enclosed in quotation marks, to identify the main features of serious games in class diagrams. The profile makes it possible to standardize the representation of several serious game models. These stereotypes can then be used to facilitate code generation to a database in order to trace and analyze game events and actions. A class diagram showing a serious game formalization with our profile is presented in Figure 2 and 3. We describe below the main stereotypes used in this diagram.

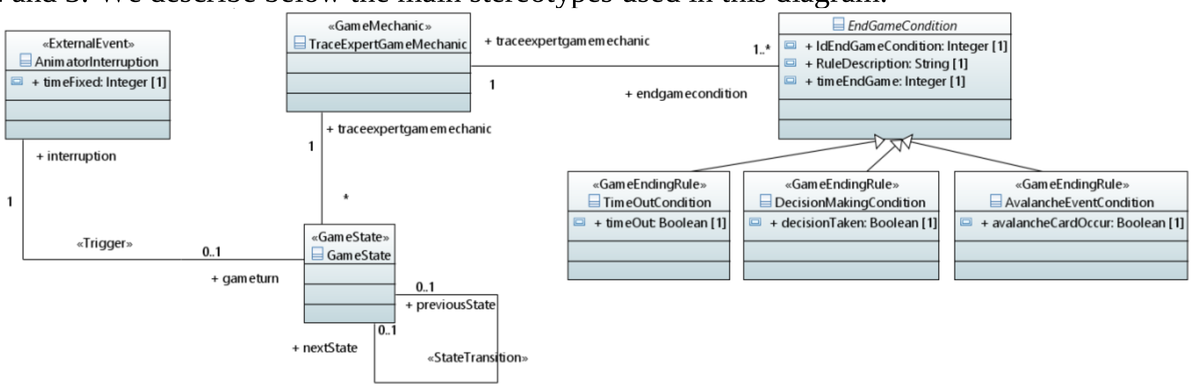


Figure 3. Example of use of our UML profile for serious game traceability modelling (part B)

A game is based on a risk assessment grid (RiskAssessmentGrid) established by a team of experts (ExpertTeam) and on several scenarios (GameScenario) inspired by real life. A game is played during a session (GameSession). Each model instance corresponds to one game session. Each session is played on the basis of a scenario (GameScenario). An action that can be done by a player (PlayerCharacter) during the game is to draw a card or to make an evaluation (GameActionAssess) on the new context (GameContext). For each turn, the game context is changed by a new information and individual assessment (each player) and team assessment (all players agree) are required (AssessmentTrace). The game is presented as a succession of turns (GameTurn) that present different states (GameState). The transition of a given state to another state is triggered either by ExternalEvent or by GameAction. GameMechanic allows describing the rules of the game such as the end of game conditions (GameEndingRule): time out reached, decision taken or fallen avalanche. The risk assessment grid contains a way of evaluation of different criteria and provides a framework for the different types of information provided to players. These criteria can be categorized (AvalancheSite or SnowWeatherConditions for example). Each of its criteria is associated with evaluation means which can be quantitative (LimitsAssessment) or qualitative (CriterionDescriptiveAssessment). For example, a condition can be evaluated as "Favorable", between 2m and 3m of snow.

CONCLUSION

In this paper, we defined the first concepts to describe the different elements to model a serious game. The goal is to propose a uniform way to represent several serious games. The proposed stereotype can be used for the modelling of other serious games. In the present paper, we use a specific case to start the test of our approach. We aim at showing the interest of such approaches. We also present the details of the serious game we developed for decision making related to avalanche risk.

In our future work, we will enrich the stereotypes and apply our proposal on several other serious games, to demonstrate the reusability of our approach. It will also be needed to propose mechanisms to generate the physical storage of game traceability data. The resulting database will facilitate the analysis of decision-making reasoning.

ACKNOWLEDGMENTS

The authors thank the colleagues involved in this project on the topic of avalanche risk assessment: Sébastien Escande and Xavier Pasquier (ANENA), François Rapin and Michaël Deschâtres (IRSTEA). The serious game development shown in this paper has been founded by DGPR (Direction Générale de la Prévention des Risque), the French Ministry of the Ecological and Solidarity Transition, in the TraceExpert project, elaborated in an agreement with the SRNH (Service des Risques Naturels et Hydrauliques).

REFERENCES

[1] J. Thompson J., B. Berbank-Green, N. Cusworth, “The computer game design course: principles, practices and techniques for the aspiring game designer”. Thames & Hudson, 2007.

[2] R. Hunicke, M. LeBlanc, R. Zubek, “MDA: A formal approach to game design and game research”, Proceedings of the AAAI Workshop on Challenges in Game AI, 1–5, 2004.

[3] D. Altunbay, M.G. Metin, M. Çetinkaya, “Model-driven Approach for Board Game Development”. First Turkish Symposium of Model-Driven Software Development (TMODELS), Ankara, Turkey, 2009.

[4] S. Tang, M. Hanneghan, “A model-driven framework to support development of serious games for game-based learning”. Developments in E-systems Engineering (DESE), 95–

100, 2010.

[5] A. Yusoff, "A conceptual framework for serious games and its validation". University of Southampton, 2010

[6] B. Marne, J. Wisdom, B. Huynh-Kim-Bang, J.M. Labat, "The six facets of serious game design: a methodology enhanced by our design pattern library". European Conference on Technology Enhanced Learning, 208–221, 2012.

[7] B.C. Ibáñez, B. Marne, J.M. Labat, "Conceptual and technical frameworks for serious games". Proceedings of the 5th European conference on games based learning, 81–87, 2011

[8] N. Aouadi, P. Pernelle, C.B. Amar, T. Carron, MDA Approach for Reusability in Serious Game and E-learning Design. International Conference on Web-Based Learning, 206–212, 2016.

[9] N. Thillainathan, J.M. Leimeister, "Serious game development for educators-A serious game logic and structure modeling language", 2014.

[10] R. Nystrom, "Game programming patterns", Genever Benning, 2014.

[11] S. Bjork, J. Holopainen, "Patterns in game design (game development series)", 2004.

[12] B. Marne, J.M. and Labat, « Implémentation de patrons de conception pour l'adaptation des parcours pédago-ludiques dans les jeux sérieux". *8ème Colloque Technologies de l'Information et de la Communication pour l'Enseignement (TICE 2012)* 69–79, 2012.

Insight into Decision Support Systems for Sustainable Geotechnical Engineering

Mrs. Pavlina Karagianni, Prof. Enrique Romero
Universitat Politècnica de Catalunya UPC
c/ Jordi Girona 1-3, Campus Nord UPC, 08034 Barcelona, Spain
pkaragianni@cimne.upc.edu, enrique.romero-morales@upc.edu
web-page: https://www.upc.edu/en

Dr. Alessandra Di Mariano, Mr. Angel Priegue
CIMNE
c/ Gran Capità s/n, Campus Nord UPC, 08034 Barcelona, Spain
alessandra.dimariano@upc.edu, cruchi@cimne.upc.edu
web-page: http://www.cimne.com/

ABSTRACT

Geotechnical engineering has a significant role in sustainability, due to its early place in the construction process and thus with a higher possibility to limit the impact of the project. Nevertheless, there is a lack of methods for evaluating geotechnical developments considering the three core pillars of sustainability. A new decision support system DSS for the sustainability assessment in the construction of geotechnical infrastructures is indeed needed to evaluate and compare geotechnical alternatives by taking into account the environmental, social and economic parameters. This paper aims at presenting a preliminary insight into a new DSS for sustainable geotechnical infrastructures and presents an approach for the development of it.

Keywords: Decision Support System, MADM Method, Construction, Geotechnical Engineering, Sustainability Assessment

Introduction

The world is encountering unprecedented challenges related to energy resources, global warming, and waste material generation. Failure to overcome these challenges will inhibit the world development and will negatively impact the standard of living and security of current and future generations [1].

The construction industry is one of the largest exploiters of natural resources, both renewable and non-renewable, which is adversely changing the environment of the earth. It depletes two-fifths of global raw stone, gravel, and sand and one-fourth of virgin wood, and consumes 40% of the global energy and 16% of water annually [2]. A vast majority of the construction industry invariably involves geotechnical engineering as one of its components that has a huge potential contribution to the project sustainable development due to its early position in the construction process. The sustainable development should implement a holistic consideration of economic growth, social development and environmental protection (Fig. 1).

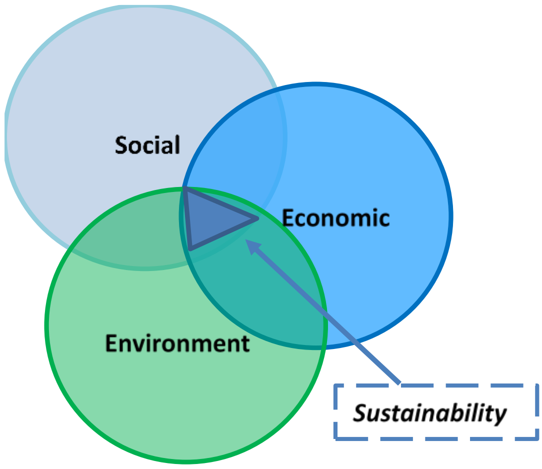


Figure 1: Sustainability Pillars

Consistent with the spirit of sustainable development, in recent years, plenty of assessment tools have been created for different categories of geotechnical projects, without though taking into account all the sustainability pillars and usually facing the lack of weighting system that might lead to misrepresentation of risk areas for certain types of projects (Fig. 2).
Geotechnical engineering is a complex area because every engineering problem can be solved in different ways, considering a variety of parameters in a usually limited amount of time. Hence, geotechnical engineers are often dealing with the pressure of decision making, which might lead to a misleading interpretation of all available options and cause problems related to sustainability targets, deadlines, standards and budgets. The absence of frameworks for the evaluation of different geotechnical techniques through the prism of sustainability is hence considered extremely important. Yet, developing a sustainability assessment and performance-based management tool for geotechnical assets needs substantial effort [3].
Those sustainability evaluation problems can be overcome by developing a decision support system DSS, based on a Multiple Criteria Decision Making MCDM method, which helps decision makers to organise and synthesise information, allowing a complex comparison between available alternatives [4, 5].

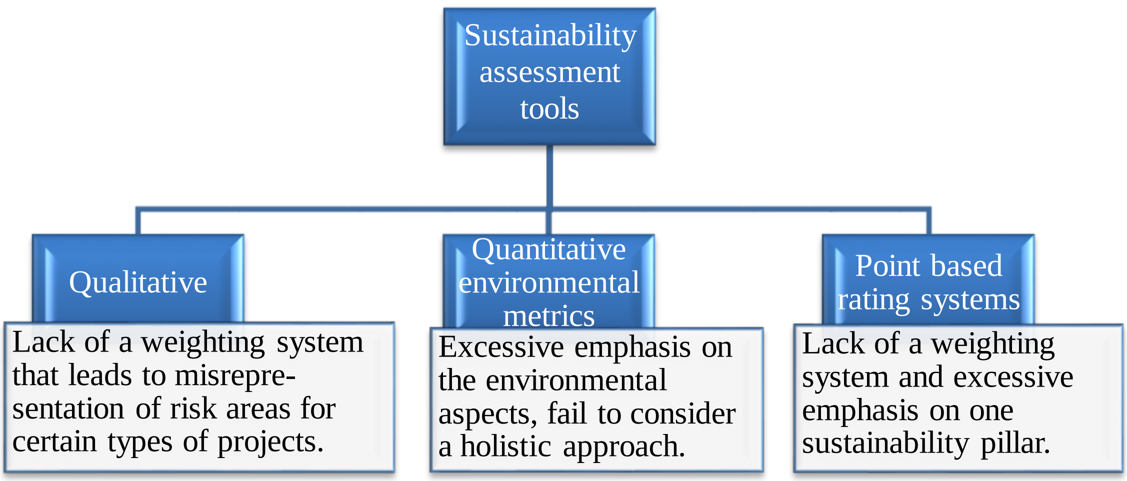


Figure 2: Existing sustainability assessment tools and encountered issues

The objective of the present research is to develop a new DSS for sustainable geotechnical infrastructures based on a MCDM method. This paper gives a first insight into the main methods that will be used for the new DSS and highlights the importance of it.

Development of a new decision support system

The MCDM method has been developed as a part of operational research, concerned with designing mathematical and computational tools approaches. It is well known in the decision-making process for supporting subjective evaluations of performance criteria. In the present work, MCDM method has been selected due to its unique characteristics such as the presence of multiple non-commensurable and conflicting criteria, possible different units among the criteria, and the presence of different alternatives [5, 6, 7].
Two different methods are usually applied to solve the MCDM problems: the Multiple Objective Decision Making MODM method and the Multiple Attribute Decision Making MADM method. MODM method deals with many objectives to come up with an optimal solution to achieve prefixed objectives. Whereas in the MADM method, the decision-making process deals with alternatives that have a variety of performance attributes and factors, either qualitative or quantitative [8]. In this case, the MADM method has been chosen to allow the comparison among alternative methods.
The MADM method is a discrete method, which generally assigns numbers to all pre-determined alternatives. It specifies how to process the attribute's information to reach an ideal choice. To model those attributes, the MADM method will be presented through a decision matrix. This matrix consists of 1) alternatives, 2) criteria and 3) relative significance of criteria. Because a wide range of criteria can be considered during a sustainability assessment, depending on the type and the specific conditions of each project, it is considered critically important to define a core set of criteria to ensure that there is some level of consistency in the assessment process.
According to previous working experiences, as well as based on available published information and literature, Table 1 contains a list of the core criteria that will be implemented in the new DSS.

Table 1: Sustainability Criteria

Sustainability Criteria		
Social	**Environmental**	**Economical**
Supply chain	Energy consumption	Material cost (removal, disposal, imported material)
Influence on surface area	Vibration pollution	Equipment cost (temporary, permanent)
Noise pollution	Waste material and usage	'On site' cost (site facilities, energy consumption, employees)
Vibration pollution	Emission of CO_2	Aid from government
Affection of existing structures and services	Environmental incidents and risk	
Interaction/distribution of current transportations and services	Usage of recycled materials	
Duration of works		

Some criteria (Table 1) are qualitative and need to be converted into quantitative scales for comparison purposes [9]. Commonly, the five-point Likert-type scales are used for this conversion. However, even though the five-point scales can convert qualitative criteria to numbers, they are not able to distinguish clearly the differences among close scores (e.g., good and very good). Because of that, a hierarchical scale consisting of 17 relative importance factors will be utilised for the comparison matrix among variable decisions.
In general, there are various MADM techniques to be used during decision-making processes. In the present research, the Technique for Order of Preference by Similarity to Ideal Solution TOPSIS method has been selected, for ranking all the available alternatives, due to its simplicity and ability to consider a non-limited number of alternatives in the decision making process. TOPSIS is a method of compensatory aggregation that compares a set of alternatives by identifying weights and normalising scores for each criterion, calculating the geometric distance between each alternative and the ideal one, which is considered the one with the best score.
Shanian and Savadago [10] noted that it is important to know the relative significance of each criterion in a MADM problem. For the determination of the weights, subjective fixed weight methods – such as the Delphi method, expert survey method, the analytic hierarchy process method (AHP), etc. – are usually used. However, those methods could eventually lead to deviations of indexes' weights due to subjective factors. On the other hand, objective-fixed weight methods, based on the inherent information of indexes to determine weights of indexes, could eliminate man-made disturbances and increase the accuracy. Shannon [11] noted that the entropy method could also be used to objectively calculate the relative weight of information, except in the cases when data needs to be quantitatively estimated. The entropy weight method is one of the objective-fixed weight methods and it is based on the information determined by the index's weights [11]. Also, Pratyush and Jian-Bo [12] believe that the entropy method can evaluate the amount of uncertainty represented by a precise probability distribution.

Due to the reasons above, in the integrated model that it is being designed, the entropy method has been chosen to calculate the relative weight of attributes and the TOPSIS method to compare a set of alternatives by normalising scores for each criterion and calculating the geometric distance between each alternative and the ideal one. As for the significance of each criterion with respect to others, the user will be able to introduce a specific value referring to a provided hierarchical scale. As a result, it is expected that DSS will provide flexibility to the user and adjustability according to the project requirements. Hence, the decision makers can be anyone involved in a geotechnical project and will be able to evaluate a project according to the core criteria.

CONCLUSIONS

It is extremely important to understand that decisions related to sustainability should not be based on the 'we have always done it that way' logic but should be adjusted and evaluated based on every project needs, by considering the multi-dimensional and multi-disciplinary aspects of sustainability.
Evaluation of geotechnical techniques by encompassing all the aspects of sustainability, from a multi-disciplinary point of view, while measuring their sustainability through each construction stage, is necessary for achieving all the sustainability credentials of a construction project. The current absence of frameworks for the evaluation of different techniques through the prism of sustainability needs to be urgently addressed.

The development of an integrated model that incorporates a DSS, can play a key role in predicting geotechnical impacts, as well as contributing to save time and reduce project costs. Such a model, on which the authors are currently working on, would help designers to select the most sustainable technique for proposed geotechnical projects, taking into account project limitations on one side and users' requirements and preferences on the other.

The new DSS, which is under development, uses the MADM method to allow the comparison of alternative methods. It normalises values and applies weighting factors with the use of the TOPSIS and entropy methods, respectively. By doing so, with the new DSS, it will be possible not only to evaluate the amount of uncertainty, through the representation of a precise probability distribution, but also to determine the geometric distance between each alternative and the ideal one.

REFERENCES

[1] R. J. Fragasz et al "Sustainable Development and Energy Geotechnology − Potential Roles for Geotechnical Engineering", KSCE Journal of Civil Engineering , 611 - 621, 2011.
[2] M.K. Dixit, J.L. Fernández-Solis, S. Lavy & C.H.Culp, "Identification of parameters for embodied energy measurement: a literature review", Energy and Buildings, 42: 1238–1247, 2010.
[3] K.L.S. Bernhardt, J.E. Loehr & D. Huaco, "Asset management framework for geotechnical infrastructure. Journal of Infrastructure Systems", ASCE, 9(3): 107–116, 2003.
[4] V. Belton & T. Stewart, "Multiple Criteria Decision Analysis: An Integrated Approach" Kluwer Academic Publishers, 2002.
[5] Mardani, A. Jusoh, E.K. Zavadskas, F. Cavallaro, & Z. Khalifah, "Sustainable and renewable energy: an overview of the application of Multiple Criteria Decision Making techniques and approaches", Sustainability 7(10): 13947– 13984, 2015b.
[6] Edmundas Kazimieras Zavadskas, J. Antuchevičienė & O. Kapliński, "Multi-criteria decision making in civil engineering: Part I – a state-of-the-art survey". Engineering Structures and Technologies , 103-113, 2015.
[7] Mardania, A. Jusoh &. E. K. Zavadskas, "Fuzzy multiple criteria decision-making techniques andapplications – Two decades review from 1994 to 2014". Expert Systems with Applications, pp. 4126 – 4148, 2015.
[8] Shanian & O. Savadogo, "TOPSIS Multiple-Criteria Decision Support Analysis for Material Selection of Metallic Bipolar Plates for Polymer Electrolyte Fuel Cell", Power Sources , 159(2), pp. 1095-1104, 2006.
[9] C.L.Hwang & K. Yoon, "Multiple Attribute Decision Making: Methods and Applications" Springer-Verlag, 1981.
[10] Shanian & O. Savadogo, "TOPSIS multiple-criteria decision support analysis for material selection of metallic bipolar plates for polymer electrolyte fuel cell". Journal of Power Sources, 1095-1104, 2006.
[11] C.E. Shannon, "A Mathematical Theory of Communication". The Bell System Technical Journal, 379–423, 623–656, 1948.
[12] Pratyush & Jian-Bo, "Multiple Criteria Decision Support in Engineering Design". London: Springer, 1988.

Risk Assessment & Security-based Approaches

Decision Support System in Thailand's Dam Safety

Apinat Dowrueng
Technical Writer
Am Ölper Berge 20, 38114 Braunschweig, Germany
apinat.dowrueng@gmail.com

Chinoros Thongthamchart
Geotechnical Engineering Research and Development Center – Kasetsart University
50 Thanon Ngam Wong Wan, Lat Yao, Chatuchak, Bangkok 10900
t_chinoros@yahoo.com
http://www.gerd.eng.ku.ac.th

Nanthiya Raphitphan, Potcharapol Brohmsubha
Electricity Generating Authority of Thailand
81 Moo 11 Bangkruai-Sainoi Rd., Sai Noi, Nonthaburi 11150, Thailand
nanthiya.r@egat.co.th, potcharapol.b@egat.co.th
http://www.egat.co.th/en/

Wanlaya Laungnarutai, Unpong Supakchukul, Arnuphap Dowrueng
National Electronics and Computer Technology Center
112 Phahonyothin Road, Khlong Luang District, 12120 Pathumthani, Thailand
wanlaya.laungnarutai@nectec.or.th, unpong.supakchukul@nectec.or.th,
arnuphap.dowrueng@nectec.or.th
https://www.nectec.or.th/en/

ABSTRACT

This paper describes the decision-making based on civil engineering expertise of the Dam Safety Remote Monitoring System (DS-RMS) initiated by the Electricity Generating Authority of Thailand (EGAT) and developed by Geotechnical Engineering Research and Development Center – Kasetsart University (GERD-KU) and National Electronics and Computer Technology Center (NECTEC). The system decides on action-based advice depending on every-day scenarios and special occurrences like earthquakes and floods. Main benefits include fast and reliable access to information on the dams for the public and being a reliever to dam executives in case of critical situations.

Keywords: decision support system, dam safety, disaster management, remote-monitoring

INTRODUCTION

The main objective of Dam Safety Remote Monitoring System (DS-RMS) [1] is to automate conventional acquisition of safety-related measurements so that comprehensive information is remotely obtained in real time. DS-RMS is specially designed to support three major dam types: concrete, embankment, and impervious faced rockfill dams. The fourteen dams, located across the country e.g. the Srinagarind Dam: a rockfill dam featuring a clay core with over 17 billion cubic meters water reservoir capacity, or the Bhumibol Dam: a concrete arch-gravity dam with over 13 billion cubic meters capacity, are equipped with reliable remote terminal units (RTUs), communication links, and computer servers installed with advanced information software, including dam behavior visualization, earthquake monitoring, flood routing simulation, and safety evaluation and warning system.

DEVELOPING DS-RMS

First of all, the Geotechnical Engineering Research and Development Center at Kasetsart University (GERD-KU) worked out evaluation methods to check the current status of the dam and issue different warnings. Following that the National Electronics and Computer Technology Center (NECTEC) developed a knowledge-based automated web application by splitting GERD’s evaluation methods into smaller flowcharts and sort them into three main zones: normal operation, earthquake events and flood events. Figure 1 shows one of these flowcharts for earthquake events. We will use this chart as an example for part of the evaluation. [2] discuses flowcharts for flood events with focus on civil engineering.

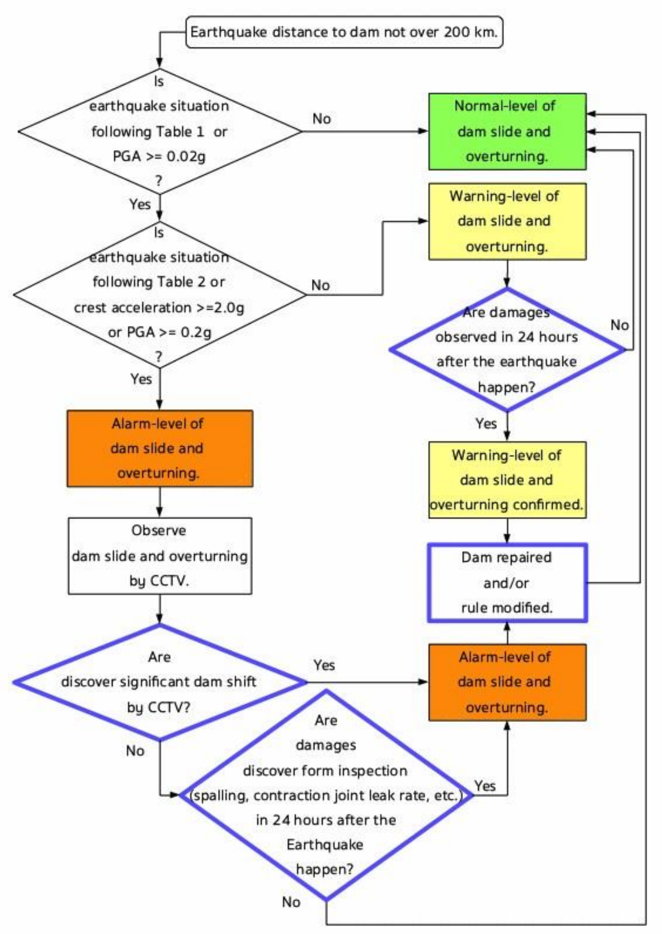


Figure 1: Example of a flowchart for earthquake events

At first the system checks whether thresholds are being exceeded as shown in the following tables and evaluates instructions to act upon or warning levels according to Figure 1.

Table 1: Earthquake Magnitude in relation to distance of earthquake epicenter to dam induced PGA more than 0.02g (acceleration on the foundation of the dam)

Earthquake Magnitude	Distance of earthquake epicenter to dam (km)
≥4	≤25
≥5	≤50
≥6	≤80
≥7	≤125
≥8	≤200

Table 2: Earthquake Magnitude in relation to distance of earthquake epicenter induced PGA more than 0.2g

Earthquake Magnitude	Distance of earthquake epicenter to dam (km)
≥ 6	≤ 12
≥ 6.5	≤ 18
≥ 7	≤ 24

The warning levels the system will use are green for Normal, yellow for Alert and orange for Alarm statuses. All the flowcharts are being evaluated separately and will output one the mentioned three colors. [3], [4] recommend two warning levels, which represent abnormal behavior of the dam, for dam surveillance as follows.

- Alert: the measured data is out of the expected range, taking into account the common changes due to cyclic or stationary loads.
- Alarm: the maximum level forecasted by the engineering board whereby the safety coefficients for the structure are surpassed.

There are also two critical differences to a typical expert system for example the system being presented in [5]:

1. DS-RMS will never decide on its own and will only issue action-based advice to dam officials, e.g. not give out an evacuation command and therefore the color red is not being used here.
2. EGAT wants to know exactly when we have an unexpected scenario that is not covered by the flowcharts. Therefore additional color: grey is defined.

DS-RMS USER INTERFACE

Summary for Executives

Picture 2 shows the summary page for executives with the identified dam safety status of all dams on the map of Thailand. If needed, a summary in text form can be opened, as well as last events of the three main zones. More detailed information on every single dam can be found by clicking on the links in the left menu panel.

Picture 2: Summary for executives

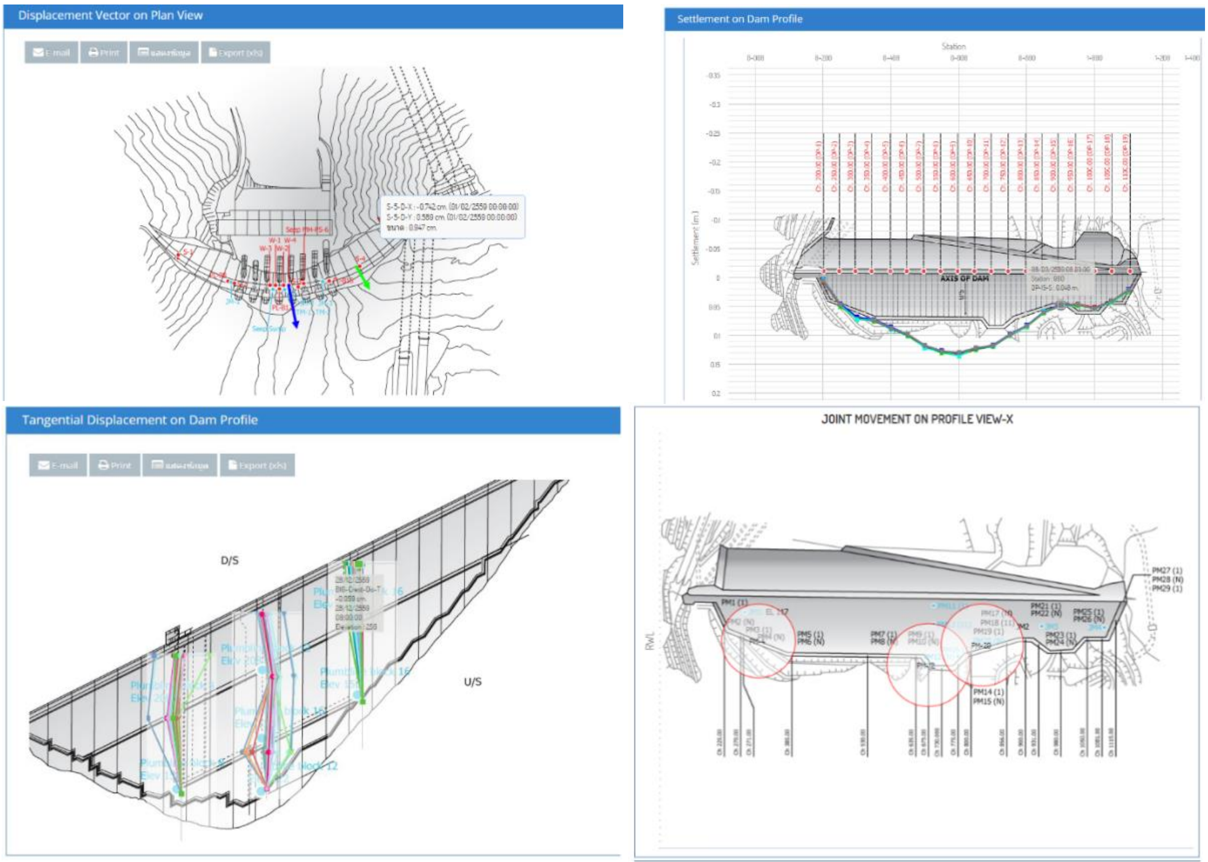
Picture 3: Detailed information for flood main zone

Detailed information on evaluation

Picture 3 shows detailed information for the flood main zone of one dam. The flood zone is split into five flowcharts for this dam type. The highest warning level for a single flowchart will be transferred to the main zone and from there to the status of the dam.

Different multidimensional graphical schematics

Dam behavior visualization was developed with features automatic data mapping mechanism and multi-dimensional displays, including cause and effect (correlation), section, plan, and profile views to ease data interpretation. The data mapping mechanism is designed to help map data of two concerned instrument types which may have different sampling rates but they are selected to be plotted together as 'cause and effect' display e.g. to find out whether piezometric head is affected by reservoir water level. The mapping mechanism searches and maps two values with the nearest timestamps by considering an acceptable time gap.



The multi-dimensional displays provide users with various data plots overlaying on graphical dam structure drawing background. Cause and effect plot shows the relationship between two variables (or two instrument types). Section/plan/profile views display the various graph types on different dam drawings where instruments or sensors are installed. For example, displacement vectors can be shown on a plan view to indicate how much and which direction the dam body moves. Settlement profile view can illustrate the settlement of the dam crest in the vertical direction. This allows EGAT operators to easily identify the problems and promptly take appropriate actions.

Picture 4: EGAT's staff by DS-RMS user training at Thailand Science Park

DS-RMS has now been in full operation since 2016. The feedback report from EGAT operators shows that the developed dam behavior visualization can significantly reduce time and effort taken in the data preparation, visualization, and analysis processes. Same as current discussions on autonomous driving, the main point is acceptance of abstract computer algorithms, which cannot be manually calculated and validated by a human: [5] and [6] or Predictive Analytics (Predictive Modeling, Machine Learning, Data Mining, etc.). The systems widespread in e.g. South Korea [7] and Australia [8] are not using Decision Support Systems as presented here.

Furthermore, the use of water resources in Thailand is very diverse: e.g. generating electricity, drinking water and agriculture whose competences are spread across multiple ministries. EGAT, responsible for power generation, is subject to the Ministry of Energy. Power distribution is responsibility of Provincial Electricity Authority (PEA: https://www.pea.co.th/en) and is subject to the Ministry of Interior which also decides on evacuations and disaster relief. The wide area flood catastrophe 2011 in Thailand shown how complex such a situation can be and that we are not quite ready yet to let computers making all decisions for us. After introducing and accepting this system, the next possible step would be to develop and test a more autonomous system in parallel to the exiting one.

CONCLUSION

The main objective here is to automate the conventional acquisition of safety-related measurement data by hand, so comprehensive information can be remotely obtained in real-time. DS-RMS is specially designed to support three major dam types: concrete, embankment, and impervious-faced rockfill dams. The fourteen dams located across Thailand are equipped with reliable remote terminal units (RTUs), communication links, and servers equipped with advanced information software, including dam behavior visualization, earthquake monitoring, flood routing simulation and a safety evaluation and warning system. The system reassures EGAT and the public nearby, especially downstream communities, that the safety of the dams is thoroughly analyzed and monitored at all times.

ACKNOWLEDGEMENT

The authors would like to thank EGAT for initializing and funding the entire DS-RMS project. C. Thongthamchart representing the Geotechnical Engineering Research and Development Center at Kasetsart University conceptualized the decision-making process behind the system with their civil engineering know-how based on [3], [4] and [9]. Finally, the National

Electronics and Computer Technology Center developed the resulting computer system containing both remote monitoring and decision support for dams.

REFERENCES

[1] NECTEC, „DS-RMS : Dam Safety Remote Monitoring System," 22 06 2016. [Online]. Available: https://www.nectec.or.th/en/innovation/service-innovation/dsrms.html. [Accessed 19 02 2019].

[2] C. Thongthamchart und P. Brohmsubha, „The Safety Criteria for Geotechnical Instruments on the Internal Erosion in Embankment Dams," in *Internation Symposium on Dams in a Global Environmental Challenges*, Bali, Indonesia, 2014.

[3] International Commission on Large Dam, „Bulletin 158 Dam Surveillance Guide," 2010.

[4] International Commission on Large Dam, „Bulletin 164 Internal Erosion of Existing Dams, Levees and Dikes, and their Foundations," 2013.

[5] M. Leone-Filho, D. A. F. Balbi, A. E. Toscano, P. S. F. Barbosa, R. M. Salgado, L. F. V. Vianna und M. N. Barreto, „Expert System for Dam Assessment and Emergency Detection," in *6th International Conference on Flood Management*, Sao Paulo, Brazil, 2014.

[6] H. Su, Z. Wen und Z. Wu, „Study on an Intelligent Inference Enginein Early-Warning System of Dam Health," *Water Resour Manage (2011)*, Bd. 25, Nr. 25, p. 545–1563, 2011.

[7] K. Y. Jung, „Safe Water Service," K water, [Online]. Available: http://english.kwater.or.kr/eng/busi/controlBusioutlinePage.do?s_mid=1190. [Accessed 17 02 2019].

[8] T. S. o. V. D. o. E. a. P. Industries, „Dams in Victoria - Safe and versatile infrastructure," 2014. [Online]. Available: https://www.water.vic.gov.au/__data/assets/pdf_file/0018/54126/Dam-Safety-Factsheet-FINALv4b.pdf. [Accessed 17 02 2019].

[9] R. Fell, P. MacGreger, D. Stapledon und G. Bell, Geotechnical Engineering of Dams, UK: Taylor & Francis, 2005.

Security risk assessment and communication

Ana Respício
LASIGE, Universidade de Lisboa
FCUL – DI, C6, Piso 3, 1749-016 Lisboa, Portugal
alrespicio@fc.ul.pt
web-page: http://respicio.at.di.fc.ul.pt/

Pedro Dias Rodrigues
EDP - ENERGIAS DE PORTUGAL SA, Digital Global Unit, SOC EDP
R. Camilo Castelo Branco, 45, 1069-227 Lisboa, Portugal
PedroDias.Rodrigues@edp.com

ABSTRACT

This talk presents the Multi-level Risk Manager, a software developed in the context of the DiSIEM EU H2020 project (http://disiem-project.eu/). The DiSIEM – Diversity enhancements for Security Information and Event Management (SIEM) – project aims to address the limitations of SIEM systems already deployed in production. Instead of proposing novel architectures for future SIEMs or modifications to existing ones, the project addresses their limitations by extending current systems, leveraging their built-in capacity for extension and customisation. These extensions have been materialised through a set of tools and components, in the form of plugins.

The Multi-level Risk Manager component relies on a hierarchical model to assess different levels of security risk in the organization. This component provides risk metrics and analytics to support decisions concerning which infrastructure configurations offer better security guarantees. In addition, it was designed to increase the capacity of security managers to communicate risk to C-level managers.

The system has been in validation and evaluation at the security operations center of EDP – Energias de Portugal SA, the Portuguese Energy producer and distributor, and one of the partners of the DiSIEM Consortium.

Keywords: Hierarchical Decision-Making, Security Risk Assessment, Risk Communication, Risk Analytics.

Methodology for a Risk Assessment Decision Support System in Vertical Farming

Francis Baumont De Oliveira, Scott Ferson
Institute for Risk & Uncertainty, University of Liverpool
L69 7ZF Liverpool, U.K.
F.baumont-de-oliveir@liverpool.ac.uk, Ferson@liverpool.ac.uk
http://www.riskinstitute.org.uk/

ABSTRACT

This paper describes a decision support system (DSS) for small-scale vertical farms (VFs) to inform investments through risk analysis. The emerging industry of vertical farming has struggled with growing pains, primarily with economic viability and proving claims of environmental sustainability. Without a robust business model, a large proportion of vertical farms fail within several years without continued investment. This issue, inherent in entrepreneurship, is magnified by high capital demand and a very steep learning curve. From scattered literature it is evident that improved knowledge transfer and scientific economic analyses for VFs are needed. A short review reveals an opportunity for a DSS, informed by lessons learned from industry and academic expertise, centralising relevant knowledge on vertical farming. It aims to assist with planning stages and early operations, whilst remaining flexible to both standard and bespoke systems. The DSS database is developed from literature using structured expert elicitation protocols. The software uses this data to stochastically simulate economic viability through first-passage time processes. The knowledge-base provides advice on common practices and contextual risk mitigation strategies. This DSS prototype facilitates risk-empowered business plans, allowing safe testing of different scenarios whilst reducing the learning curve for VFs.

Keywords: Vertical Farming, Urban Agriculture, Risk Assessment, Business Sustainability, Uncertainty Quantification, Decision Support

INTRODUCTION

Background

Feeding a predicted 9.8 billion people by 2050 [1] on a planet stressed by climate change, water scarcity, soil degradation, ageing rural populations, and rising levels of urbanisation, will require constant innovation in resilient farming methods. In this paper, vertical farming is defined as the practice of hydroponically cultivating crops indoors in vertically stacked layers or inclined surfaces. Modern vertical farms (VFs) utilise indoor farming techniques to take advantage of controlled-environment agriculture (CEA) technology within repurposed shipping containers, warehouses, purpose-built plant factories, greenhouses, and other structures. With CEA processes and proper management, all environmental factors can be controlled (including light spectra) for optimal growing conditions called growing recipes. The most common crops are leafy greens, salads, herbs, microgreens, vine crops and some fruits and flowers. Hops and commodity crops are more difficult due to their high energy demands. Many advantages can be realised when compared to conventional agricultural methods. These include:

1. Minimising horizontal space requirements, with increased yield per unit area [2];
2. Reducing dependence on pesticides or herbicides [3];
3. Cutting water consumption drastically by roughly 90% [4];
4. Producing reliable year-round crop, in soil-less environments independent of weather [5];
5. Reducing the necessity for storage, transport and refrigeration by local production [5];
6. Increasing food safety due to reduced variables of wildlife and increased traceability [5];
7. Operating electrically to reduce direct dependence on fossil fuels [5];

Growing food in cities, near the point of consumption, is a rational strategy. Similarly, fresh high-value plants for pharmaceuticals or packaged or prepared food can be cultivated close to the point of manufacture. One study in 2010 even estimates $80-$160 billion in ecosystem services could be realised annually from global and intense urban agriculture adoption [6].

In the past decade, the industry has experienced huge growth, primarily due to the reduction in operational expenditure (OpEx) and capital expenditure (CapEx) from advancements in light-emitting diode (LED) technology, automation and sophisticated greenhouse technology [8,9]. Industry market reports claim that the vertical farming market has risen from $400 million in 2013 to $1.5 billion in 2016 [10], expecting it to reach between $6.4 and $13 billion by 2024 using compound annual growth rates [10–12]. The rising demand for local fresh produce by environmentally conscious consumers has fostered this growth [7]. However, a lack of available trained workforce, still-high CapEx, and troubles with profitability have continued to plague the industry [8, 9].

Challenge of business sustainability

Despite the many perceived benefits over conventional agriculture, the vertical farming industry has struggled to be widely adopted. It is rightly challenged with scepticism due to its limited crop viability, high energy demands, high CapEx demand, and its financial uncertainty. Its economic viability has been identified as the largest obstacle to realising VFs, and whilst it has been said to be profitable [9], the learning curve is steep, and the financial risk is high. Naturally, due to the time and investment required, there is secrecy around business models and lessons learned [10]. However, some organisations are realising collaboration will be crucial for its success [10,11].

Preliminary results indicate approximately 85% of food-focused VFs fail within several years unless given further capital investment [12]. These failures are more acute because of the high CapEx investments. Additionally, VFs require a complex “urban food-water-energy nexus” approach, requiring cross collaboration among researchers in business, academia and government policy analysis not commonly seen in urban planning and design [13]. There are many complex variables to be managed, and it is essential to mitigate costly mistakes early in a project.

Paul Gauthier, a researcher from Princeton University summarises well: “Vertical farms might work as a technical concept. Thriving as a business that transforms agriculture is another matter.” [10]. What can be done to reduce the barriers to entry and ensure sustainable growth of an incipient industry still finding its feet? Two clear needs are bridging the knowledge between various sectors and conducting a robust economic analysis for VFs [14,15].

Currently, detailed financial analysis of CapEx, OpEx and revenues have been hard to produce, due to the unprecedented nature of combining architecture, agriculture, and other specialist areas [15]. Calculations tend to be for particular scenario, and they are difficult to generalise. Projects struggle to realise an acceptable return on investment (ROI) above 10%.for investors.

This paper proposes a decision support system (DSS), acting as a hub for compiling lessons learned with an adaptable economic model library to produce robust risk assessment under uncertain scenarios. The model would utilise data collected from literature, experts and focused interviews about VFs (existing and shuttered). By using imprecise data analysis techniques, the model can assess viability of a VF in particular scenarios. This viability assessment would inform decisions for business plans (location, product choice, product recovery, etc.) with a risk mitigation strategy tailored to the user’s needs.

REVIEW OF RELEVANT SYSTEMS

Hydroponics and greenhouses

There is a wide pool of DSS literature for greenhouse management which could be collated and adapted for small-scale VFs. The management of production in a greenhouse is similar to indoor vertical farms and requires decision making on many tasks and time scales [16].

Primarily decisions are related to management of crop growth conditions, irrigation, propagation, and the controls of environment based growth models [16–20]. A DSS has been developed formalising expert practice and scientific knowledge to generate set-point values for greenhouse cultivation resulting in energy savings of 5-20% [21]. Another DSS is able to quantify energy costs based on different control strategies [22]. Greenhouse growers use prediction tools for disease, yields, production planning, pest management, and cost-benefit analyses [16,23–25].

Vertical farming systems

There are some studies and tools that exist for calculating costs and return on investments for VFs [14,26,27]. However, they rely upon deterministic calculations and there is no tool that utilises stochastic methods to model uncertainty. Without such a supported financial case, developers and investors will struggle to commit to the development of a VF. A flexible economic estimation software “VFer” was developed by researchers at the University of Nottingham, as a reliable way to overcome lack of detailed designs or historical data [15]. This knowledge-based software helps with preliminary designs based on some initial decisions and produces a cost and revenue report. One general conclusion from VFer is that researchers must work with existing small-scale pioneers of vertical farming to develop and refine increasingly accurate cost models [15].

Some companies and suppliers of VF systems also use in-house software for deterministic projections of ROI and yields. In a limited sample (n=5), every VF system supplier agreed with the idea that a stochastic model (incorporating the learning curve of a VF) would be preferable to aid decision-making and more accurate forecasting [28].

Lastly, MIT’s Open Agricultural Initiative builds open resources to enable acceleration of digital agricultural innovation. They collaborate with many sectors and are developing an open-source ecosystem of technologies and datasets, sharing crop growth recipes developed through machine learning [29]. The inaccessibility of crop growth recipes has been identified as a challenge in the industry and would be centralised in the proposed DSS.

METHODOLOGY

Structure of the decision support system

The DSS facilitates better decision-making, utilising a database, a knowledge-base, a model library, and a user interface. Clearly defining these components and how they interact is important in order to avoid an overly complicated system. A poorly designed software may result in complex interfaces, unnecessary development and time-consuming simulations which often hinder their use [30]. Figure 1. shows the system structure of the proposed DSS, which has been adapted from “VFer” to incorporate several steps, improve accuracy and include risks and uncertainties [15].

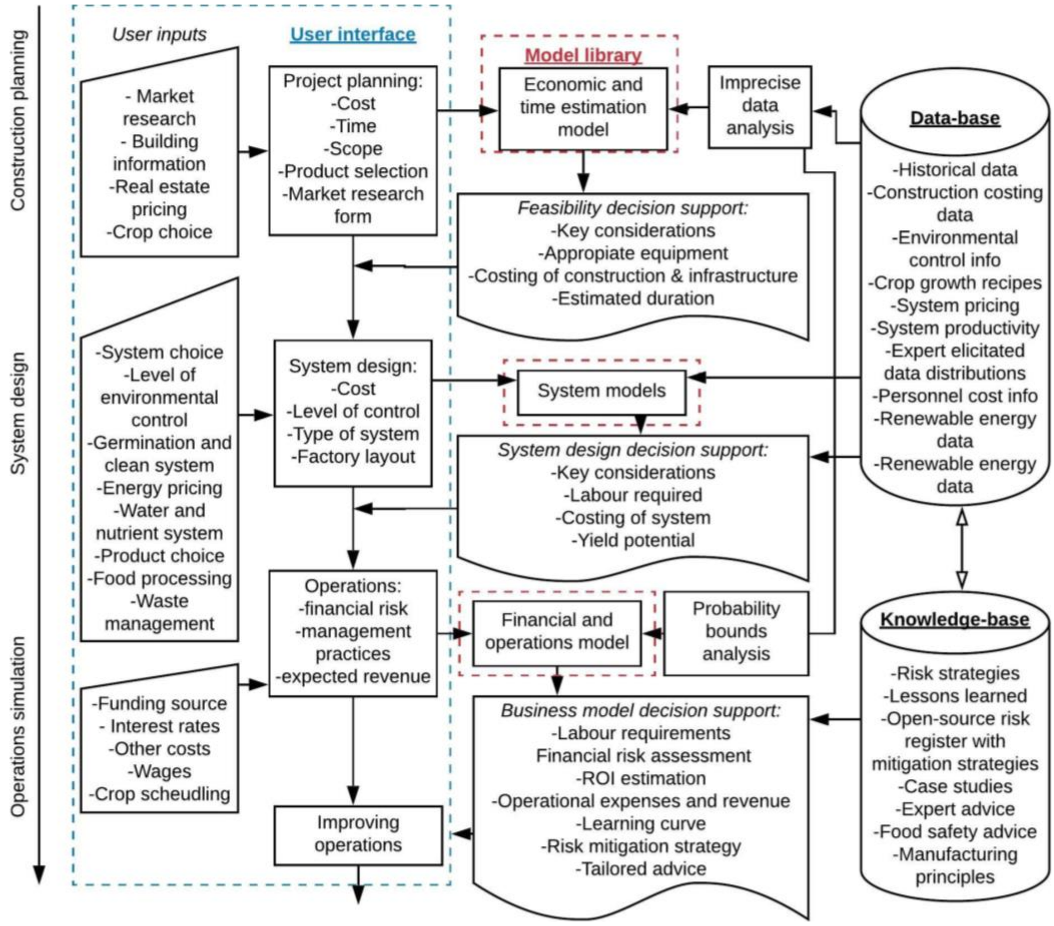


Figure 5 System processing structure for DSS

The database – The database is where necessary data are input and stored in the DSS. Data is sourced from available literature and studies conducted on operational and shuttered VFs, both in industry and academia. The VF studies include data gleaned from interviews using the structured expert elicitation protocol 'IDEA' [31], to mitigate contextual biases and improve accuracy and transparency of resulting judgements. IDEA includes the four-step elicitation procedure, which is important for decisions in which empirical data are absent or insufficient [31]. The analysis of collected data is used to establish uncertainty bounds that estimate distributions of time to reach peak operational performance of VFs (representing the learning curve), fluctuations in yield at peak performance, and other important considerations. Historical data from case studies informs estimates for duration, costs and labour of developing VF projects. The database would also include technical information about crops, using the OpenAg initiative for instance, construction costing data, location data, system costing from suppliers, and data related to environmental control.

The model library – To achieve economic forecasts for a VF, an element of risk and uncertainty should be incorporated. Models with random parameters are called stochastic models [32]. The bounds established in the database, are used for probability bound analysis (expressed as bounds on cumulative distribution functions called "p-boxes"), which enable risk calculations without requiring over-precise assumptions about parameter values or distribution shape [33]. Due to the scarcity of data, imprecise probability techniques such as probability bounds analysis [33] will then be applied prior to being used in the model library. This method is computationally faster than Monte Carlo and is guaranteed to bound the correct answer, it often produces optimal solutions [33]. This is a crucial step in order to reduce computing time and cost. Computing with p-boxes allows modelling with significant uncertainty, most importantly in calculations of risk in business sustainability. P-boxes can be used to model the event of bankruptcy after crossing a threshold defined as the first-passage time, used commonly in economics [34]. A financial model which utilises this approach is used to assess the financial risk. Figure 2 shows a financial risk assessment of a VF, its p-box primarily falls within moderate risk category with some substantial risk until the five-year mark. The risk categories are defined as follows (labelled by the red Xs in Fig. 2):

- *Critical:* 50% probability of bankruptcy within 3 years
- *Substantial risk:* 25% probability of bankruptcy within 5 years
- *Moderate risk:* 10% probability of bankruptcy within 10 years
- *Safe:* Less than 10% probability of bankruptcy within 10 years

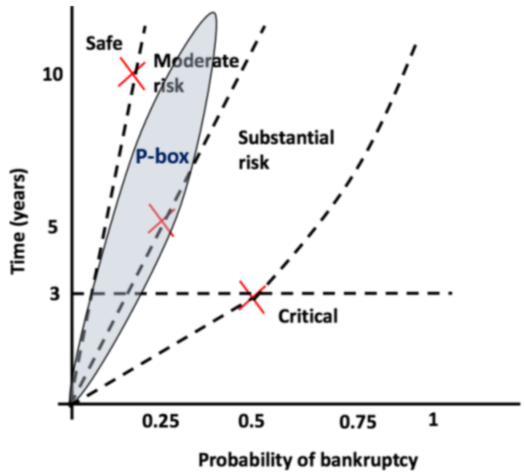


Figure 6: Risk assessment graph for the probability of bankruptcy adapted from [35]

The knowledge-base – The knowledge-base in the DSS works in combination with the database to guide the user through best-practices, applying lessons learned, and providing reports which help build a risk-empowered business plan. The knowledge provided enables a user to think about risk trade-offs before capital investment and through user-decisions identify the top threats to a VF project based on sensitivity analysis and case-studies. A built-in risk-register would be developed from conducted experiments and observations (examples can be found here [9]) and can be used to acquire strategies to deal with common problems. Assigned probabilities would be subjective regarding the likelihood and impact given a prior, however would avoid vertical farmers searching through scattered literature for solutions.

The user interface – The user interface enables the user to interact with the DSS. The system takes users through a series of menus, allowing them to choose to start from planning stages to operational stages. A user-experience design approach was adopted because it is paramount to aid quasi-real-time decision making. As the user goes through the stages (illustrated in Figure 1), it will support the development of their business plan from planning to operations. Figure 3. shows a mock-up of a menu for farm characteristics, allowing them to select an appropriate business model (such as high-value: low volume, speciality crops, see [26] for more examples) based upon a built-in market research form.

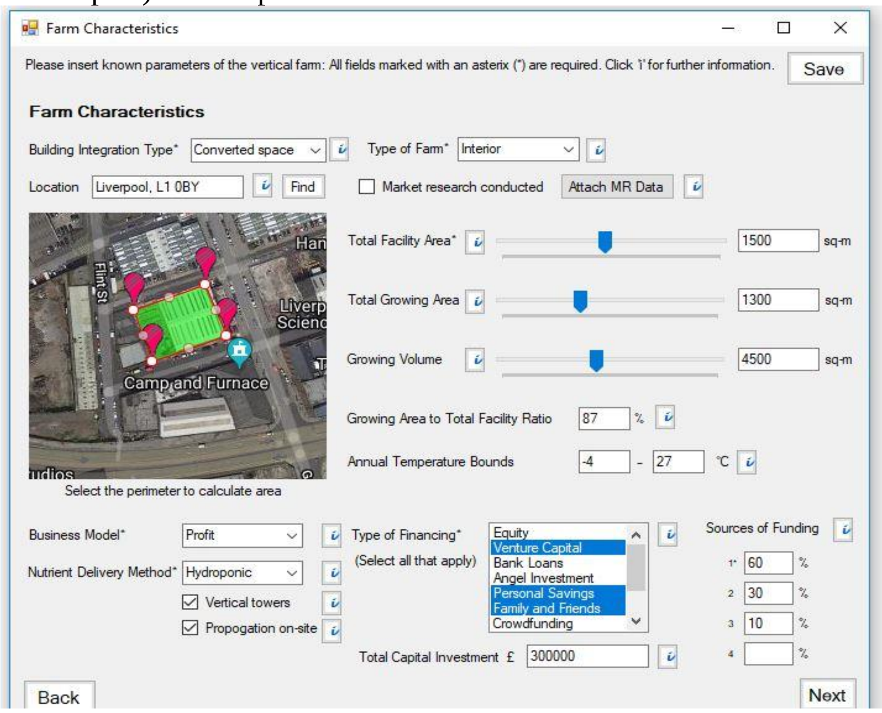


Figure 7 A mock-up of the farm characteristics menu of the DSS, Map data from ©2019 Google [36]

CONCLUSIONS

It is claimed by industry leaders that economic viability of VFs is possible with a robust business model [9,26]. However, these claims require comparative and economic data to have scientific validity. A large obstacle to profitability of this industry is knowledge acquisition on how to design and run an efficient VF business. Risk-empowered business plans and technological improvements should drive higher margins and reduced operational overheads. From the literature, more robust economic analysis of VFs is identified as needed [14,15]. A DSS is proposed that incorporates the risks and uncertainty of these complex systems and can facilitate more accurate business models with financial estimation and risk assessment. The interviews and implementation of the DSS are currently underway.

Acknowledgements

This work would not be possible without support from Low Carbon Eco-Innovatory (LCEI), Farm Urban and my PhD supervisor Dr Ronald Dyer.

REFERENCES

[1] United Nations Department of Economic and Social Affairs Population Division, "World Population Prospects The 2017 Revision Key Findings and Advance Tables," 2017.

[2] D. Touliatos, I. C. Dodd, and M. Mcainsh, "Vertical Farming Increases Lettuce Yield per Unit Area Compared to Conventional Horizontal Hydroponics," *Food and Energy Security,* vol. 5, no. 3, pp. 184–191, 2016.

[3] P. Marks, "Vertical Farms Sprouting All Over the World," *New Scientist*, no. January 2014, pp. 3–5, 2014.

[4] G. L. Barbosa *et al.*, "Comparison of Land, Water, and Energy Requirements of Lettuce Grown Using Hydroponic vs. Conventional Agricultural Methods," *International Journal of Environmental Research and Public Health,* vol. 12, no. 6, pp. 6879–6891, 2015.

[5] D. Despommier, *The Vertical Farm: Feeding the World in the 21st Century*. 2010.

[6] N. Clinton *et al.*, "A Global Geospatial Ecosystem Services Estimate of Urban Agriculture," *Earth's Future*, 2018.

[7] P&S, "Global Vertical Farming Market Size, Share, Development, Growth and Demand Forecast to 2023 - Industry Insights," 2017. [Online]. Available: https://www.psmarketresearch.com/market-analysis/vertical-farming-market. [Accessed: 01-Jun-2018].

[8] K. Benke and B. Tomkins, "Future Food-Production Systems: Vertical Farming and Controlled-Environment Agriculture," *Sustainability: Science, Practice and Policy*, vol. 13, no. 1, pp. 13–26, 2017.

[9] M. Liotta, P. Hardej, and M. Nasseri, "An Examination of Shuttered Vertical Farm Facilities," in *Aglanta Conference Panel*, 2017.

[10] C. Kummer, "Trouble in the High-Rise Hothouse," *NEO.LIFE*, Dec-2018. [Online]. Available: https://medium.com/neodotlife/vertical-farming-paul-gauthier-76e81ace79d0 [Accessed: 27-Feb-2019]

[11] L. Burwood-Taylor, "Japan's Indoor Ag Sector is Becoming More Collaborative," *AgFunder*, 2018. [Online]. Available: https://agfundernews.com/japans-indoor-ag-sector-is-becoming-more-collaborative.html. [Accessed: 25-Jun-2018].

[12] P. Gauthier, "Personal Correspondance via Email with Researcher at Princeton University." 2018.

[13] JPI Urban Europe, "Sustainable Urbanisation Global Initiative (SUGI)/Food-Water-Energy Nexus," 2018. [Online]. Available: https://jpi-urbaneurope.eu/calls/sugi/. [Accessed: 27-Feb-2019].

[14] C. Banerjee and L. Adenaeuer, "Up, Up and Away! The Economics of Vertical Farming," *Journal of Agricultural Studies*, vol. 2, no. 1, p. 40, 2014.

[15] Y. Shao, T. Heath, and Y. Zhu, "Developing an Economic Estimation System for Vertical Farms," *International Journal of Agricultural and Environmental Information Systems*, vol. 7, no. 2, pp. 26–51, 2016.

[16] R. R. Shamshiri *et al.*, "Advances in Greenhouse Automation and Controlled Environment Agriculture: A Transition to Plant Factories and Urban Agriculture," *International Journal of Agricultural and Biological Engineering*, vol. 11, no. 1, pp. 1–22, 2018.

[17] N. D. Clarke *et al.*, "Development of the Harrow Greenhouse Manager: A Decision-Support System for Greenhouse Cucumber and Tomato," *Computers and Electronics in Agriculture*, vol. 24, no. 3, pp. 195–204, 1999.

[18] M. K. Gupta, D. V. K. Samuel, and N. P. S. Sirohi, "Decision Support System for Greenhouse Seedling Production," *Computers and Electronics in Agriculture*, vol. 73, no. 2, pp. 133–145, 2010.

[19] J. A. Sánchez-Molina, N. Pérez, F. Rodríguez, J. L. Guzmán, and J. C. López, "Support System for Decision Making in the Management of the Greenhouse Environmental based on Growth Model for Sweet Pepper," *Agricultural Systems*, vol. 139, pp. 144–152, 2015.

[20] A. Pawlowski, J. A. Sánchez-Molina, J. L. Guzmán, F. Rodríguez, and S. Dormido, "Evaluation of Event-Based Irrigation System Control Scheme for Tomato Crops in Greenhouses," *Agricultural Water Management*, vol. 183, pp. 16–25, 2017.

[21] M. Tchamitchian, R. Martin-Clouaire, J. Lagier, B. Jeannequin, and S. Mercier, "A Daily Set Point Determination Software for Glasshouse Tomato Production," *Computers and Electronics in Agriculture*, vol. 50, no. 1, pp. 25–47, 2006.

[22] O. Körner and G. Van Straten, "Decision Support for Dynamic Greenhouse Climate Control Strategies," *Computers and Electronics in Agriculture*, vol. 60, no. 1, pp. 18–30, 2008.

[23] F. J. Baptista, "Modelling the Climate in Unheated Tomato Greenhouses and Predicting Botrytis Cinerea Infection," Universidade de Evora, 2007.

[24] G. Dimokas, M. Tchamitchian, and C. Kittas, "Calibration and Validation of a Biological Model to Simulate the Development and Production of Tomatoes in Mediterranean Greenhouses during Winter Period," *Biosystems Engineering*, vol. 103, no. 2, pp. 217–227, 2009.

[25] P. R. Fisher, R. D. Heinqa, N. Ehler, J. Heinrich, and P. Karlser, "A Decision-Support System for Real-time Management of Easter Lily. Scheduling and Validation," *Agricultural Systems*, vol. 54, no. I, 1997.

[26] S. Hughes, "Vertical Farming: Does the Economic Model Work?", Nuffield Farming Scholarships Trust, 2018.

[27] J. Eaves and S. Eaves, "Comparing the Profitability of a Greenhouse to a Vertical Farm in Quebec," *Canadian Journal of Agricultural Economics.*, vol. 66, no. 1, pp. 43–54, 2018.

[28] Personal correspondance via email in 2018 with vertical farming system suppliers.

[29] MIT, "MIT - OpenAg Open Phenome Project," 2018. [Online]. Available: https://www.media.mit.edu/projects/open-phenome-project/overview/. [Accessed: 01-Mar-2019].

[30] D. Zhang, X. Chen, and H. Yao, "Development of a Prototype Web-Based Decision Support System for Watershed Management," *Water*, vol. 7, no. 2, pp. 780–793, 2015.

[31] V. Hemming, M. A. Burgman, A. M. Hanea, M. F. McBride, and B. C. Wintle, "A Practical Guide to Structured Expert Elicitation Using the IDEA Protocol," *Methods in Ecology and Evolution*, vol. 9, no. 1, pp. 169–180, 2018.

[32] M. A. Burgman, S. Ferson, and H. R. Akcakaya, *Population and Community Biology Series 12: Risk Assessment In Conservation Biology*, First. Chapman & Hall, 1993.

[33] P. Van den Brink *et al.*, "Bounding Uncertainty Analyses," *Application of Uncertainty Analysis to Ecological Risks of Pesticides*, no. March 2016, pp. 89–122, 2010.

[34] S. Redner and J. R. Dorfman, "A Guide to First-Passage Processes," *American Journal of Physics*, vol. 70, no. 11, pp. 1166–1166, 2002.

[35] H. R. Akcakaya, "Population Viability Analysis and Risk Assessment," in *Proceedings of Wildlife 2001*, 1992.

[36] Google (n.d.), Map of Liverpool, United Kingdom. *Google Maps,* 3 March 2019 [Online] Available: https://maps.google.com

Classification of Integrated Security Areas (ISA) according to the Crimes Severity: A Machine Learning and Big Data Approach

Jean Gomes Turet, Ana Paula Cabral Seixas Costa
Universidade Federal de Pernambuco
Management Engineering Department, Av. da Arquitetura, S/N, 50740-550, Recife-PE, Brazil
jeanturet@gmail.com; apcabral@hotmail.com

ABSTRACT

Over the past few years, a growing concern has taken place among authorities over crimes committed worldwide. It has not been different in Brazil. High crime rates have encouraged government authorities involved in public safety to identify solutions to minimize crimes considered VCA (Violent Crime Against Assets). In this context, one alternative to both plan and manage security is in the division of ISA (Integrated Security Areas) neighborhoods. However, the local government needs to identify the degree of crime severity of each ISA in order to establish prevention policies and, therefore, resources management, which will make it possible for the Secretariat for Social Defense and other entities involved to establish sets of actions intending to minimize the crime rate. Furthermore, it is evident that changes are constantly happening in the crime behavior tendencies in each ISA, requiring a different set of daily actions from authorities. Therefore, the purpose of this study is to analyze the crime behavior in the region under study based on data provided by the website "onde fui roubado" and classify each ISA according to the severity of the crime through algorithms of classification of machine learning. Data analysis will take place from the KDD methodology and Big Data tools will be used to handle the large volume of data. Thus, this paper presents a framework that will allow the identification of the crimes severity in each ISA based on Big Data and Machine Learning.

Keywords: public safety, big data, machine learning, knowledge discovery database

INTRODUCTION

During the last years, public security has been the subject of special care by several organs and countries, especially in Brazil, where the number of crimes and violence against the population is increasing. These crimes are classified as Violent Crime against Assets (VCA), which includes all types of robberies, and the Violent Intentional Lethal Crime (VILC), which include homicides, robberies and injuries followed by death [1]. Thus, government agencies have invested in technologies to minimize these crimes and, consequently, offer the population more quality of life and safety.

In this context, there has been an increase in investments on information technology (hardware) infrastructure, but a smaller investment in intelligence, such as data mining, machine learning and the analysis and storage of large volumes of data (Big data). These technologies, when linked to a knowledge extraction methodology (KDD), can provide important results for analysis and prediction of crimes in the region.

The KDD process is composed of a set of steps in order to obtain new knowledge about a given domain from a raw database. It can be said that the KDD process comprises the whole cycle that the data travels through until it changes. The KDD has five stages: data selection, data preprocessing, data transformation, data mining and data interpretation, knowledge [2][3].

Specifically, in the region under study, these tools can be important allies in the prevention of crime, due to the fact that in Brazil one of the ways of administering the occurrence of crimes, as well as their combat strategies, is the division of neighborhoods from the classification of Integrated Sacurity Areas (ISA). The ISA are basic territorial units for the purpose of planning and executing programs, actions and police operations, setting goals and finding crimes (referencing). In this way, it is up to each ISA to supervise and administer the security of the region in which they are located [4].

However, in order to combat crime, the local government needs to identify the degree of crime severity of each ISA so they can establish prevention policies and, therefore, resource management, which will make it possible for the Secretariat for Social Defense and other entities involved to establish sets of actions intending to minimize crime. Furthermore, it is evident that changes are constantly happening in the crime behavior tendencies in each ISA, requiring a different set of daily actions from authorities

Thus, this work aims to perform the analysis of crime behavior in the region under study from the data provided by the website “onde fui roubado” and classify each ISA (Integrated Security Areas) according to the severity of the crime through algorithms of machine learning classification. Data analysis will take place from the KDD methodology and Big Data tools will be used to support the large volume of data.

RELATED STUDIES

In the literature, there are several studies that deal with public safety. These researches bring different contexts, such as large data analysis where Big Data becomes an important tool, whereas others seek to perform data mining to search for patterns and, consequently, extract knowledge. Studies that deal with machine learning and establish models of prediction and classification to aid decision-making processes in public safety are also common.

Specifically, the studies involving Big Data, Data Mining and knowledge extraction, the [5] develops a model that can predict criminal behaviors from a spatial choice with data mining techniques. [6] provides a framework that can identify crimes from different criminal activities. This is an investigative study on using spatiotemporally tagged tweets to predict crime. [8] extends this prediction by estimating Twitter and Kernel density and includes information coming from Twitter (the social media network) within an intelligent DSS.

On the other hand, in the public safety area, there are only a few studies that associate this issue with machine learning. The existing researches deal with it in a somewhat more general way, such as [7], in which this study builds a decision support system for predictive policing. The authors establish a DSS where there is a prediction model that aims to establish a correct geographical distribution of police patrol based on data provided by the Spanish National Police Corps.

These works, as well as others presented in the literature show important gaps that must be studied and analyzed. First, it is verified that works presented in the literature on public security are limited to the construction of decision support systems taking into account historical data and generic data mining algorithms. It is important to integrate these tools for more satisfactory results.

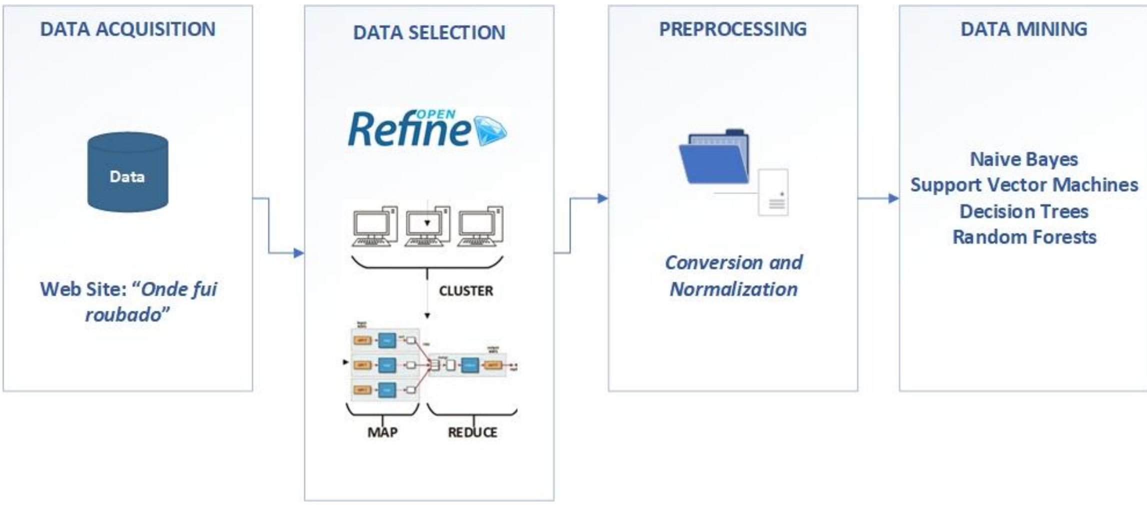
Second, there are no works that incorporate Big Data tools (Haddop Framework) with the KDD (Knowledge Discovery Database) and machine learning process. It is verified the importance of integrating these tools due to the fact there are large amounts of data that need to be processed in order to the results from the machine learning algorithms to be satisfactory.

Third, a large part of the works presented that involving classification in machine learning is associated with other themes, such as classifying students in universities; classification of smokers according to historical data. In this context, this work brings a differential, due to the fact the study is focused on the classification of Integrated Security Areas that constantly need to make decisions in real time taking into account the crime of locality, as well as the degree of crime severity. This model is the beginning of a study where models of prediction and crime prevention will be incorporated in the region under study.

FRAMEWORK PROPOSED

In order to simplify some steps of the KDD process, this study establishes a standard framework for analyzing the data coming from the website "onde fui roubado" available in "ondefuiroubado.com.br". This framework is divided in four main parts (Figure 1):

Figure 1: Framework Proposed



The model is only to classify the ISA taking into account the crime severity. Based on that, this framework can be integrated with other model to prevent the crime, but this paper only discusses the classification considering the classification machine learning algorithms (Naïve Bayes, Support Vector Machines, Decision Trees and Random Forests).

The first step is to acquire the data. The data comes from the website "onde fui roubado". This website aims to register violent crimes against property from citizens who suffer data crimes. In this context, the site provides hundreds of contents about occurrences, divided into several criteria, such as: Age, Type of Crime, Most Stolen Objects, Location of Crimes, among others. Thus, it is necessary at this stage to extract and store the data.

After this step, it is important to perform the data refinement by using algorithms, such as Open Refine, and its analysis based on Big Data. For this purpose, the Hadoop Framework was used to analyze large volumes of data in a practical and efficient way, based on two processes: map and reduce [9]. This step was performed on a high-powered computer, with 32 gb of ram memory and the latest generation I5 core processor.

For this data to be in accordance with the standards of machine learning and WEKA software, there is a need for data conversion and normalization. In this case, the data are standardized for later analysis. There are tools in the python language available that can be used for the normalization process, since performing a large amount of data and normalization manually would take more time. Normalization is divided into three main points: relation, attributes and data. In the first part, it is necessary to identify the name of the application. After that, one must identify the attributes that will be analyzed and then associate with the data. Each data will consist of previously stipulated attributes

Finally, the last step is a machine learning/data mining process. In this case, it is possible to use some algorithms available for the prediction process and, consequently, classification. For the last part, the Weka tool can be used. This tool enables the execution of different data mining algorithms as well as machine learning. For this work, the algorithm J4.8 was used for the construction of the decision tree. This algorithm was chosen because of its versatile characteristics, which allow the classification of records without the need to discretize numerical values, such as age, in the generated data sets.

The algorithm J4.8 is an open-source Java implementation of the C4.5 algorithm, which is based on a strategy that expands a decision tree by locally choosing which attributes to use to partition the data [6].

RESULTS

After the data coming from the website "onde fui roubado" was defined, it was necessary to perform their refinement. There is a total of five ISA, so the data is refined and directed to each ISA based on the neighborhoods that each one is related to. The refinement of these data is performed using the Open Refine algorithm.

On the other hand, it was verified that there were data available from the last five years (2014-2019), thus requiring Big Data tools. The data were analyzed from the Hadoop framework and counted on computers of last generation.

After data extraction, refinement and analysis, the need for data normalization was verified. As previously reported, the WEKA software standard requires the data to be divided into three main parts. In this context, the necessary attributes were identified, such as sex, age, type of crime and stolen objects. Type of crime is subclassified into seven categories: Assault Attempt (TA), Theft (F), Vehicle Break-in (AV), Group Assault (GA), Robbery (R) and Vehicle Theft (RV). In relation to stolen objects, they are subdivided into seven categories: Cellphone (C), Wallet (CA), Money (DI), Documents (DI), Purse or Backpack (BM) and Credit Card).

After the normalization of the data, the next step of the framework is to execute the data classification algorithm. As previously reported, the J 4.8 algorithm will be used for the classification of ISA according to its type of crime, taking into account that the studied city holds five ISA (Figure 2).

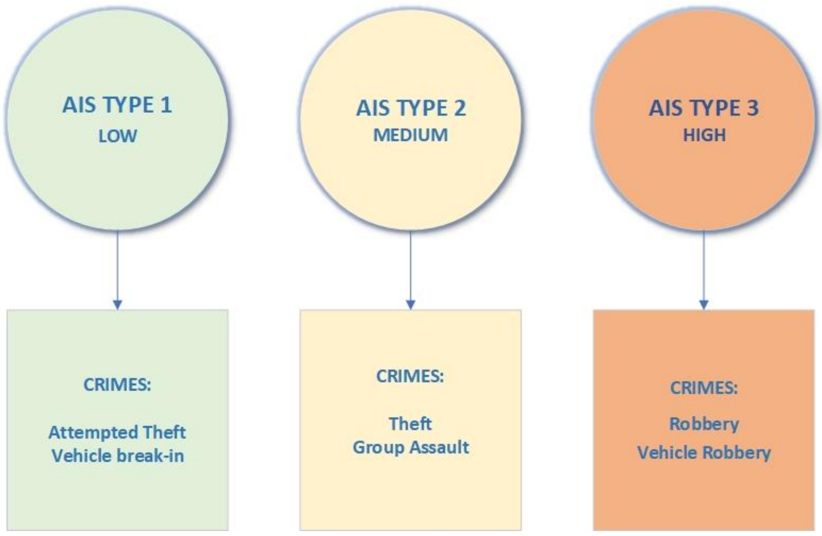


Figure 2: ISA Crime Severity Classification

The types of crimes were defined according to severity, being divided into three main points. ISA with low crime severity are those in which most crimes are attempted and there is no enforcement. ISA with crimes of medium severity are those in which there are robberies and assaults to groups. On the other hand, ISA with high severity crimes are those in which there are mostly robberies and vehicle theft.

There are several possible combinations to run the classification algorithm in the WEKA tool. The one used by this work is the Use training Set, which allows the loaded dataset to be used to train the algorithm and generate a classification model. With this, it will be possible to predict and classify ISA according to the severity of crimes.

In this context, for each set of data of each ISA was used the algorithm J 4.8 in the Mode Use training Set and as final result (Table 1)

Table 1: Final Result

ISA	CLASSIFICATION	CRIME SEVERITY
ISA 01	ISA TYPE 2	MEDIUM
ISA 02	ISA TYPE 2	MEDIUM
ISA 03	ISA TYPE 1	LOW
ISA 04	ISA TYPE 3	HIGH
ISA 05	ISA TYPE 1	LOW

It is identified that among the five ISA analyzed, the algorithm classified two of them as in low criminal severity level, two categorized as medium and one as high crime. It should be noted that the top rate is 84.5% in this mode (algorithm J 4.8, type use training set). This makes it possible to classify these integrated areas of security to establish more efficient actions. On the other hand, this type of algorithm becomes important since there are constant changes in the types of crimes and, consequently, changes in the ISA classifications, allowing real-time analysis and more efficient actions to combat crime.

CONCLUSIONS

The purpose of this study was to analyze crime behavior in the region of study based on the data provided by the website where I was “onde fui roubado” and to classify each AIS according to the severity of the crime through classification algorithms of machine learning. In this context, a framework was established to support this process based on KDD.

As a result, it is concluded that it is possible to predict and classify ISA according to the severity of the crime based on some parameters, such as the age of the victim, neighborhood where the crime occurred, type of crimes and stolen objects. In this way, the organs involved in public security have an important mechanism to try and reduce crime in these places from the public security policy more efficient and planned, since they will understand the crime behavior that each AIS is related.

As for limitations, this work only evaluated the VCA (Violent Crime Against Assets) and it could be interesting to extend to other types of crimes, as is the case of Violent Intentional Lethal Crime (VILC). Using other classification algorithms as well as other modes of analysis becomes interesting for increased assertiveness and comparative effect.

ACKNOWLEDGEMENTS

This work was partially funded by CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) and CNPQ (Conselho Nacional de Desenvolvimento Científico e Tecnológico) in Brazil

REFERENCES

[1] Social Defense Secretary . Available in: <http://www.sds.pe.gov.br/images/indicadores/CVLI/CVLI_MENSAL_POR_REGI%C3%83O.pdf.> Access in: 11 jan, 2019

[2] U. Fayyad and R. Uthurusamy, "Evolving data mining into solutions for insights.," Communications of the ACM, vol. 45, pp. 28-31, 2002.

[3] P. R. M. Cereda and J. Jose, "Adaptive data mining: preliminary studies," in IEEE Latin America Transactions, vol. 12, no. 7, pp. 1258-1270, Oct. 2014.

[4] Social Defense Secretary . Available in: < http://www.sds.pe.gov.br/.> Access in: 11 jan, 2019

[5] Xue, Y; Brown, D.E: Spatial analysis with preference specification of latent decision makers for criminal event prediction. Decision Support Systems. 560-573 (2006)

[6] Li, S-T; Kuo, S-C; Tsai, F-C. An intelligent decision-support model using FSOM and rule extraction for crime prevention. Expert Systems with Application, 7108-7119, 2010.

[7] Camacho-Collados, M; Liberatore, F. A Decision Support System for predictive police patrolling, Decision Support Systems 75 (2015) 27-37.

[8] M.S. Gerber; Predicting crime using Twitter and Kernel Density Estimation, Decision Support Systems 61 (1) (2014) 115-125

[9] White, T. Hadoop: The Definitive Guide, 4th Edition, O'Reilly Media, 2015.

Special Talks

30 Years of Decision Support: Where did it all go wrong, and where did it all go right?

Philip Powell
University of Groningen, Netherlands
Executive Dean, Pro Vice-Master (Enterprise and Innovation) and Professor of Management at Birkbeck, University of London

This session takes a thirty year retrospective on decision support research, and the impact of that research in practice. It asks where it all went wrong over these last three decades, but balances this by assessing progress too. While we were wrong to worry about lots of things 30 years ago – data provision, processing, model building etc – we were right to concern ourselves with other things such as the lack of impact on practice. We did not worry about a host of things then – ethics, equality, privacy etc – that have arisen as our access to data has exploded. Should we have been better at worrying or not worrying?

I was lucky enough to be a founder member of the EURO Working Group on Decision Support Systems EWG-DSS and to observe, and sometimes participate in, the developments that followed. Now it is time for a look back, and for some thoughts about how we should look forward.

Monitoring & Engineering Methods for DSS

Time-aware Knowledge Graphs for Decision Making in the Building Industry

Herwig Zeiner, Wolfgang Weiss, Roland Unterberger, Dietmar Maurer
JOANNEUM RESEARCH Forschungsgesellschaft mbH, Austria,
http://www.joanneum.at

Robert Jöbstl
Haas Fertigbau GmbH, Austria, http://www.Haas-Group.com

ABSTRACT

The phenomenal progress in the development of the Internet of Things (IoT) has already had tremendous impact on almost all industrial sectors and, finally, on our everyday life. The ongoing total digital transformation leads to entirely new time-aware services and creates new and more pro-active business opportunities.Knowledge graphs are becoming more and more popular in different domains, and also for the integration of sensor networks. In this project we make knowledge graphs time-aware through a set of general temporal properties relevant for the integration of sensing networks. Time-aware knowledge graphs enable us to do time series analysis, find temporal dependencies between events, and implement time-aware applications. The requirements for the temporal properties derive from a use case of residential real estate, with the aim to enable the occupants to interact with their houses.

Keywords: Building Industry, Time-aware Analytics, Knowledge Graph, Decision Making.

Find and cite the full work as:

Zeiner, H., Weiss, W., Unterberger, R., Maurer, D., Jöbstl, R., 2019. Time-aware Knowledge Graphs for Decision Making in the Building Industry, in: Freitas, P.S.A., Dargam, F., Moreno, J.M. (Eds.), Decision Support Systems IX: DSS Developments & Future Trends, Lecture Notes in Business Information Processing. Springer International Publishing, Cham, pp. 52-64.

An Agricultural Prototype DSS LANDS for monitoring the main crop productions in Sardinia

Gianni Fenu, Francesca Maridina Malloci
University of Cagliari
V.Ospedale 72, 09124, Cagliari, Italy
fenu@unica.it, francescam.malloci@unica.it

ABSTRACT

DSS have evolved significantly since their early development in the 1970s. They have been applied in different sectors as business, financial, medical and many others. Recently their application is strongly increasing in the agricultural sector due to continuous climate changes and the need to conduct more productive and sustainable agriculture. In this context, DSSs have become notable tools to enhance the agricultural production process assisting farmers in the decision-making process in choosing the economically, socially or environmentally best alternative solution. In this paper, we describe the prototype agricultural DSS developed in collaboration with Laore Sardinia Agency for monitoring the main crop productions in Sardinia. The DSS collects, organizes, integrates and analyses several types of data with different mathematical models. In particular, the DSS is being tested to forecast risk of late blight of potato, one of the most devastating diseases world over, including Sardinia.

Keywords: Decision Support System, Decision-Making, Data Analysis, Precision Farming.

INTRODUCTION

- Decision Support Systems have become notable tools to enhance the agricultural production. Agricultural production is highly dependent on weather, climate and water availability and is adversely affected by the weather and climate-related disasters [1]. Natural disasters can result in complex issues related to crop production. It is not always possible to prevent the occurrence of these natural events, but a proper planning can considerably reduce their effects. So far farmers have made in-season decisions based on their experiences and intuition. Nevertheless, their experiences are insufficient to predict a decision-making process for a long term, which can improve yield productivity and avoid unnecessary cost related to harvesting, use of pesticide and fertilizers. In addition, by 2050, according to the Food and Agriculture Organization (FAO), the climate changes are expected to cause water scarcity and serious declines yield of the most important crops in developing countries. It means that the agriculture process will have to adapt to climate change, but it can also help mitigate the effects of climate change through the recent technologies as Decision Support Systems (*DSS*).

A DSS can be defined as a computer-based system that supports decision makers in solving a decision problem. [2] These tools can lead users through clear steps and suggest optimal decision paths or may act more as information sources to improve the evidence base for decisions [3]. Recently, they have been introduced in the agriculture as an indispensable tool to face the growing challenge of conducting sustainable agriculture which increase the quantity and quality of agricultural output while using less input (water, energy, fertilizers, pesticides, etc.). This new modern farm approach that bases its applicability on the use of technologies to detect and decide what is "right" is called *Precision Farming* (PF) [4, 5]. Nowadays, many governments in the world are investing big amount of money to encourage the researchers and companies to develop Decision Support Systems which use agricultural data to help the adoption of Precision Farming.

In this paper, we describe the system and the tests conducted through the DSS LANDS (*Laore Architecture Network Development for Sardinia*) developed. It supports Sardinian farmers in decision-making and it manages different data in order to forecast and increase yield productivity and decrease the costs of agricultural operations. The DSS, it has been developed in collaboration with the *Laore Sardinia Agency*. *Laore Sardinia Agency* deals with providing advisory, education, training and assistance services in the regional agricultural sector.

The paper is structured as follows. Section II provides background information and outlines the reasons that drove the adoption and no-adoption of DSSs in Europe, especially in Italy. Section III describes the architecture and the forecasting models used in the short case study. We conclude the paper in section IV.

DECISION SUPPORT SYSTEMS

DSS have evolved significantly since their early development in the 1970s. Over the past three decades, DSS have taken on both a narrower or broader definition, while other systems have emerged to assist specific types of decision-makers faced with specific kinds of problems [2]. One of the first definitions was given by Keen and Morton [6] that defined decision support systems as computer systems that collect resources and use the ability of computer to increase quality of decisions by focusing on semi structured problems. Recently, a DSS is defined as human-computer systems which collect information, process information and provide information based on computer systems [7]. However, the researchers agree that the main objective of DSSs is to support and improve decision making [8, 9, 10].

DSS can be composed of four main subsystem which are Data Management subsystem, Model Management subsystem, Knowledge-based subsystems and User Interface subsystem [8]. The functionality of Data Management subsystem is to manage the data that will be used as information to make decisions in the Knowledge-based subsystem. The Model component consists of a variety of models that assist decision makers in decision making. The Knowledge-based is the hearth of the system and it manages the problem-solving process to generate the final solution. The User Interface allows the users to encourage the interaction with the system to obtain information.

Generally, DSS has been classified into three categories based on problems for decision making: structured, unstructured and semi-structured.

Agricultural Decision Support Systems

DSSs have been introduced in the agriculture as an indispensable tool mainly for two reasons. First, to face the continuous climate changes that cause serious damage to production. Second, to conduct a more sustainable agriculture which increase the quantity and quality of agricultural production while using less water, energy, fertilizers and pesticides, or rather, to support the Precision Farming technologies.

In the last decade, their applications have increased thanks to the advent of new technologies, such as Cloud Computing, Data Mining, Machine Learning, Artificial Intelligent and major investments by numerous research agencies and governments all over the world.

Agricultural DSSs perform the following activities: (i) they collect, organize, and integrate several types of information required for producing a crop; (ii) they analyse and interpret the information; and (iii) they use the analysis to recommend the most appropriate action or action choices. For example, DSSs can provide farmers information on plant growth or plant disease risk useful for scheduling treatments according to the actual need of the plant. [11]. However, designing a DSS is quite complex; it requires knowledge from various multidisciplinary areas, such as crop agronomy, computer hardware and software, mathematics and statistics to analyse data. For example, to understand crop growth, it is necessary to know how each variable affects crop growth [12].

In a global level, in the agricultural sector, there is not a single agricultural DSS adopted worldwide, but over the years several DSSs have been developed for a wide range of cultivation practices concerning crop management and crop irrigation. Many of them have been developed and evaluated with different crops and different climatic conditions.

A review [13] identified five fields of applications: Diagnostic-Forecasting DSSs, Advisory DSSs, Control DSSs, Educational – Informational DSSs, Operational DSSs. Although the use of DSS simplifies decision-making in agricultural production and it is applied in several application sectors, DSSs have not been adopted with great enthusiasm by managers of farms. Their adoption has been low. Many researches have been conducted for understanding the reason of DSSs non-adoption in agriculture. These researches identified the following factors that influence the adoption of DSSs by farmers: profitability, user-friendly design, time requirement for DSS usage, credibility, adaptation of the DSS to the farm situation, information update, and level of knowledge of the user [14]. However, many of these factors, have been reduced by the increased availability of personal computers, increased access to the Internet, and increased development of web-based systems [15] The adoption and the development of agricultural DSSs in Europe was faster than in Italy. The factors that have limited its diffusion have been identified in [16] that recognize as the main cause the difficulty of using precision technologies in a heterogeneous territory with particular characteristics.

In the Europe context, the literature [17] identifies the most relevant DSSs from two thousand to today: DSSAT, APSIM, CropSyst, EPIC and STICS. The *decision support system for agrotechnology transfer* (DSSAT) is a collection of independent programs that operate together. It incorporates models of 16 different crops with software that facilities the evaluation and application of the crop models for different purpose [18]. The *Agricultural Production Systems Simulator* (APSIM) contains an array of modules for simulating growth, development and yield of crops, pastures and forests and their interactions with the soil. It has been used in a broad range of applications, including support for on-farm decision making, farming systems design for production or resource management objectives, assessment of the value of seasonal climate forecasting [19]. The *cropping systems simulation model* (CropSyst) CropSyst is a suite of programs designed to work co-operatively, providing users with a set of tools to analyse the productivity and the environmental impact of crop rotations and cropping systems management at various temporal and spatial scales [20].The *Environment Policy Integrated Climate* (EPIC) is able to manage decisions related to drainage, irrigation, water efficiency, erosion (wind and water), atmospheric conditions, fertilizer, the control of pests, sowing dates, tillage and waste management cultivation [21] The *Simulateur mulTIdisciplinaire pour les Cultures Standard* (STICS) simulates crop growth as well as soil water and nitrogen balances driven by daily climatic data. It calculates both agricultural variables (yield, input consumption) and environmental variables (water and nitrogen losses) [22].

In spite of the European context several DSSs have been adopted since their appearance in the agricultural sector, in Italy few DSSs have emerged to provide decision support systems. Recently, their adoption is intensifying thanks to increase in the use of Precision Farming technologies. The diffusion of these technologies has been slow due to the following factors: heterogeneous environments, territorial characteristics, age/level of education and company size [16]. To Incentivise employment and scientific research is the Ministry of Food and Forestry Agricultural Policies, which in [16] emphasizes the importance of developing specific tools for data analysis, with DSS functions to tackle the ongoing climate changes that are compromising the main crops of the territory. Since today, in Italy have emerged DSSs for crop management, mainly for wine and cereal production and irrigation management. Analysing the literature, among the major contributions emerge *Vite.net* for the decision-making support of the vineyard, *Granoduro.net* for decision support durum wheat crop and IRRINET for decision support for irrigation. *Vite.net* is developed for sustainable management of vineyards and is intended for the vineyard manager. The system provides in real-time several information for each vineyard as the defence against fungal disease and insects, the growth of the plant, the thermal and water stresses and many others [11]. *Granoduro.net* provides plot-specific and up-to-date decision supports about weather, fertilisation, crop growth, weed control, and disease and mycotoxin risk [23]. IRRINET system provides to farmers a day-by-day information on how much and when to irrigate crops, implementing a real-time irrigation scheduling. The latter is also used in Sardinia [24].

The contribution of this paper is the development of an agricultural DSS for monitoring the main crops in Sardinia, where the DSSs adoption have been slow due to the conformation and heterogeneity of the territory that requires the development of specific decision support systems.

DSS LANDS PROJECT

DSS LANDS was developed to help Laore technical and Sardinian farmers in decision-making about agricultural management based on the principles of Precision Farming. It was designed mainly to take data-driven decision and not to replace the decision maker.

The goals of LANDS are to: (i) optimize the resources management through reduction of certain inputs (e.g., chemicals and naturals resources, etc.) (ii) predict crop risk situations (e.g., diseases, weather alerts etc.) (iii) increase the quality of decisions for field management (iv) reduce environmental impact and production cost. It integrates different and specific modules for monitoring the main crop productions in Sardinia: citrus, artichoke, wheat, corn, olive, potato, peach, tomato, rice, vine. Currently, the DSS proposed is a prototype being tested for monitoring the potatoes crop.

ARCHITECTURE

The agricultural DSS is composed of three components:

(i) an *integrated system* for semi-real-time monitoring of crop components and storage of their data; These sources include ARPAS (*Regional Agency for the Protection of the Sardinian Environment*) weather stations, field sensors and external providers.

(ii) a *models system* which performs through several mathematical and forecasting models a cross and dynamic analysis of different types of data. Their elaboration and interpretation allow us to provide the best strategies to be applied in the field in order to forecast possible risk event situations which can damage the production.

(iii) a *cross-platform application* used by Laore technical and farmers to upload crop data collected during the field survey and to visualize the up-to- date information for managing the cultivation in the form of alerts and decision supports. It is availabled by smartphone, tablet and personal computers with different operating systems. These features allow the farmers to take advantage of the application without worrying about the device in use, to access it in any place (e.g., in the field, in the company etc.) and to simplify and enhance the agricultural management process. All information is in a graphic format that uses symbols and colours to advice and inform in an immediate, effective and unambiguous way the status of each crop management component. Internet connectivity also allows a timely updating of the features as soon as new analysis results are available and without any user intervention.

The Fig. 1 describes a conceptual diagram of the system with three main stages. In the first stage the data are collected at fixed intervals from different sources: weather stations, external providers and data uploaded to the cross-platform by Laore technicals during the field survey. In the second stage, the data are received from the *Data Receiver* which manages and controls the quality of data and then it stores them into *Env DB* (*Environmental Database*) and *Potato DB*. After that, the data are analysed through several agricultural mathematical models.

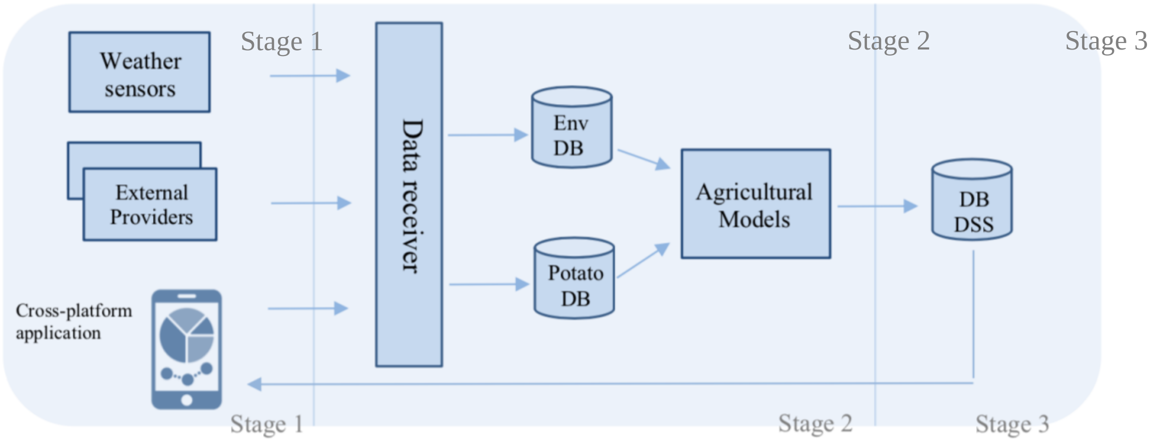


Figure 1. Conceptual diagram of the system

Finally, in the third stage, the output is stored and sent to the cross-platform application for the interpretation by the decision maker. The output is visualized in the application as graphs and guidelines through different and specific dashboards. Each dashboards are a collection of widgets that give to the farmer an overview of the metrics and let them monitor many metrics at once, so they can quickly check the health of their cultivation.

Case of Study

LANDS was tested during the 2018 spring season to forecast and tackle the risk of *Phytophthora Infestans* cryptogamic attacks for potato crop also known as *late blight* or *potato blight*. Potato blight is one of the most devastating diseases of potato world over, including Sardinia. In the Region the continuous climate changes such as the rains close together, the high humidity and the abrupt changes of the temperatures are putting at risk the potatoes production. For this reason, the experimentation phase started as a support in the decision-making process of this cultivation.

The tested are conducted in the potato fields monitored and managed by the Laore Agency.

We have implemented two disease prediction models retrieved from literature: *Negative Prognosis model* and *Fry model* [25, 26]. The joint use of the two algorithms allows to forecast the period which it is opportune to carry out fungicide treatments useful against the appearance of the pathogen.

The models identify the number of treatments need during a growing season as a function of time and meteorological data acquired continuously from ARPAS weather stations.

The analysis of the *Negative Prognosis Model* predicts the period where the late blight epidemics are not likely to occur and the timing of the first treatment. In order to achieve an accurate prediction, the system receives, manages and stores with fixed frequency the following data: hourly temperature of the day, hourly humidity of the day, hourly winds, day degrees calculated with different methodologies, Eto calculated with different mathematical formulas. From these data, the model takes as input: hourly temperature (°C), relative humidity (%), and rainfall (mm).

After the server has received the input parameters the model calculates with different formulas the risk values and the accumulated risk values. This last, is the values that allows to determinate the date of the first treatment. The Fig. 2 shows the trend of the accumulated risk index recorded from 12/03/2018 to 29/04/2018.

The tested conducted allowed to identify a local threshold which recognize when the disease is expected to occur. The warning period is indicated when the accumulated risk value exceeds the threshold of 130 and the first treatment is applied when the threshold reaches the value 150. In the case of Figure 3 the first treatment was carried out 13/04/2018.

To estimate the treatments after the first we developed the *Fry model*. The model calculates the spraying intervals based on the *blight units* and *fungicide units*. Blight units are calculated according to the number of consecutive hours that relative humidity is greater than or equal to 90%, and average temperature falls within any of six ranges (< 3, 3-7, 8-12, 13-22, 23-27 and >27 C). Fungicide units are calculated based on daily rainfall (mm) and time since last fungicide application. Decision rules about when fungicide should be applied are generated based on cumulative blight units or fungicide units since last spray.

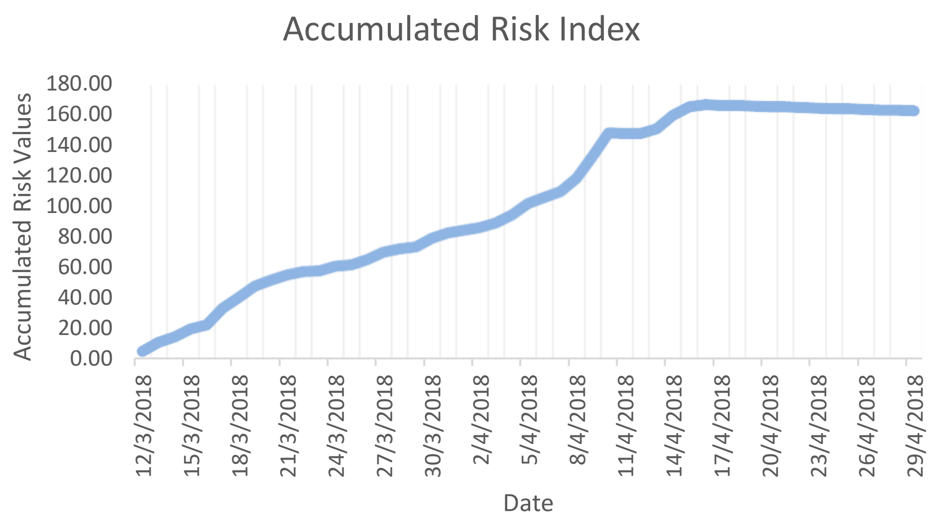


Figure 2: Accumulated risk recorded from 12/03/2018 to 29/04/2018

The experiments carried out allowed to outline the best criteria for local conditions through the *Fry model* developed. The treatments after the first are indicated when one of the following cases occurs: (i) the accumulated precipitations are greater than 20 mm, (ii) the risk value of the previous night is 8 and also the sum of the blight units exceeds 40 for cultivar susceptible.

CONCLUSIONS

In the present paper we have seen how the DSSs are widely used in the agricultural sector. They have become notable and indispensable tools to conduct a more sustainable and productive agriculture which is difficult to sustain due to the continuous climate changes.

Although several DSSs for monitoring various cultures have been developed, their adoption has been slow for two reasons: technical limitations of the DSSs and to farmer attitude towards DSSs. Today, the situation is changing thanks to the increased availability of personal computers, increased access to the Internet and increased development of web-based systems. Even in Italy and especially in Sardinia few DSS have been adopted. The major contribution of this work is the development of the DSS LANDS in collaboration with the Laore Sardinia Agency to monitor the main crops in Sardinia, a place where the adoption/diffusion of DSS is complicated for the territory heterogeneity. Currently, the DSS is a prototype being tested for monitoring the potato culture. In particular, the DSS through the *Negative Prognosis Model* and the *Fry Model* elaborates weather data from meteorological stations to forecast the period in which is opportune to carry out fungicidal treatments against the pathogen late blight outbreak. The short case of study conducted allowed to adapt, calibrate and outline the local parameters in order to produce accurate predictions.

However, LANDS is at an early stage of the project. To date, it is still early to be able to assess the benefits of its use in the field. Future experiments will allow to validate predictive dynamical models and evaluate if LANDS is the tool able to respond to the challenges emerging in the agricultural field according to Precision Farming methods.

REFERENCES

[1] M. V. K. Sivakumar, M. Ray, H.P. Das, "Natural disasters and extreme events in agriculture: Impacts and mitigation", January 2005.

[2] J.P.Shim, M. Warkentin, J. F.Courtney, D. J.Power, R. Sharda, C. Carlsson, "Past, present, and future of decision support technology", Decision Support System, V.33, pp.111-126, June 2002.

[3] D. C. Rose, W. J. Sutherland, C. Parker, M. Lobley, M. Winter, C. Morris, S. Twining, C. Foulkes, T. Amano, L. V. Dicks, "Decision support tools for agriculture: Towards effective design and delivery", Agricultural System, V.149, pp.165-174, November 2016.

[4] R. D. Grisso, M. M. Alley, P. McClellan, D. E. Brann, and S. J. Donohue, "Precision Farming. A Comprehensive Approach", 2009

[5] Zhang, N., Wang, M., Wang, N., "Precision agriculture—a worldwide overview. ", Comput. Electron. Agric., V. 36, pp. 113–132, November 2002.

[6] P. G. W. Keen, M. S. S. Morton, "Decision support systems: An organizational perspective", MA, USA: Addison-Wesley, V. 35., 1978.

[7] Y. Sheng, S. Zhang "Analysis of problems and trends of decision support systems development", E-Business and Information System Security (EBISS), Wuhan, 2009.

[8] E.Turban, J. E. Aronson, T. P Liang, " Decision support systems and intelligent systems (7th ed.)" , New Jersey, USA.: Pearson Prentice Hall, 2005.

[9] G. Fenu, M. Nitti, "Strategies to carry and forward packets in VANET", Communications in Computer and Information Science, 166 CCIS (PART 1), pp. 662-674, 2011

[10] L. Boratto, S. Carta, G. Fenu, R. Saia, "Using neural word embeddings to model user behavior and detect user segments", *Knowledge-Based Systems*, 108, pp. 5-14, 2016

[11] V. Rossi, F. Salinari, S. Poni, T. Caffi, T. Bettati, "Addressing the implementation problem in agricultural decision support systems: the example of vite.net®", Computers and Electronics in Agriculture,2014.

[12] P. Taechatanasat, L Armstrong, "Decision Support System Data for Farmer Decision Making", Proceedings of Asian Federation for Information Technology in Agriculture, 2014

[13] B. Manos, A. Ciani, Th. Bournaris, I. Vassiliadou, J. Papathanasiou, "A taxonomic survey of decision support systems in agriculture". Agricultural Economics Review 5, 2004.

[14] D. Kerr, "Factors influencing the development and adoption of knowledge based decision support systems for small, owner-operated rural business", Artificial Intelligence Review 22 (2), 2004.

[15] V.P Jones, J.F. Brunner, G.G. Grove, B. Petit, G.V. Tangren, W.E. Jones, "A web-based decision support system to enhance IPM programs in Washington tree fruit", Pest Management Science, 2010.

[16] Mipaff, «Linee Guida per lo Sviluppo dell'Agricoltura di Precisione in Italia,» 2017. [Online]. Available: https://www.politicheagricole.it/.

[17] D. Holzworth, V. Snow, S. Janssen, I. N. Athanasiadis, M. Donatelli, G. Hoogenboom e J. W. T. White, "Agricultural production systems modelling and software: Current status and future prospects", Environmental Modelling & Software, 2014.

[18] J. Jones, G. Hoogenboom, C. Porter, K. Boote, W. Batchelor, L. A. Hunt, "The DSSAT cropping system model", European Journal of Agronomy, V. 18, p.p. 235-265, 2003.

[19] B. Keating, P. Carberry, G. Hammer, M. Probert e M. Robertson, "An overview of APSIM, a model designed for farming systems simulation", European Journal of Agronomy, V. 18, p.p. 267-288, 2003.

[20] C. Stockle, M. Donatelli, R. Nelson, "CropSyst, a cropping systems simulation model", European Journal of Agronomy, V. 18, p.p. 289-307, January 2003.

[21] R. Izaurralde, J. Williams, W. McGill, N. Rosenberg e M. Jakas, "Simulating soil C dynamics with EPIC: model description and testing against", Ecological Modelling, vol. 192, n. 362-384, 2006

[22] N. Brisson, C. Gary, E. Justes, R. Roche, B. Mary, D. Ripoche, D. Zimmer e J. Sierra, "An overview of the crop model stics", European Journal of Agronomy, vol. 18, pp. 309-332, 2003.

[23] V. Rossi, V. Manstretta, M. Ruggeri, "A multicomponent decision support system to manage Fusarium head blight and mycotoxins in durum wheat", World Mycotoxin Journal, 2015.

[24] P. Mannini, R. Genovesi, T. Letterio, "IRRINET: Large Scale DSS Application for On-farm Irrigation Scheduling", Procedia Environmental Sciences, vol. 19, 2013, pp. 823-829, 2013.

[25] W. E. Fry, A. E. Apple, J. A. Bruhn, "Evaluation of potato late blight forecasts modified to incorporate host resistance and fungicide weathering", Phytopathology, 1983.

[26] J. A. Bruhn, W. E. Fry,"Analysis of potato late blight epidemiology by simulation modelling", Phytopathology, 1981

Ranking the Challenges of Hotel 4.0: Priorities according to the Experience of Implementing Innovative Water and Electricity Monitoring Technologies

Javad Jassbi[1,2*], Mohammad Hossein Ghodsirad[3], Koorosh Aslansefat[1], Mohammad Julashokri[1] and José Barata[1]

[1] Center of Technology and Systems (CTS), UNINOVA

[2] Department of Mechanical and Industrial Engineering, Faculty of Sciences and Technology, Nova University of Lisbon, 2829-516 Caparica, Portugal

[3] Optishower, Temptation Keeper Lda., Santa Cruz, Madeira, Portugal
* j.jassbi@uninova.pt

ABSTRACT

Industry 4.0, known as "the Fourth Industrial Revolution", has been blowing among different industries and changing the paradigms of doing businesses. Although all industries are under the influence of emerging technologies, the rate of success and cause of failure of transition processes are still not clear, and interest researchers. The challenges of technology transfer, technology management, and technology implementation are not new, but the fourth revolution has its own characteristics and dilemmas in practice and Hotel 4.0 is not an exception. This paper, discusses the challenges of implementing innovative Water and Electricity Monitoring Technology in Hotel industry, and the final ranking model helps understand the priorities of the challenges. The challenges could be seen in both service provider and client sides. Moreover, the main challenges were prioritized and ranked using Analytical Hierarchy Process (AHP) technique. Although the level of consistency, questions the accuracy of the weights, but the ranking of challenges could show the main concerns and importance of potential dilemmas which should be taken into account. This could help have a better understanding of future challenges, decide for strategies to improve the system, and increase the readiness. The work is based on interviews and brainstorming sessions with a group of technicians involved in the projects and information elicited from real cases in Europe.

Keywords: Hotel 4.0, Innovative Monitoring Technologies, I4.0 Readiness, Analytical Hier-archy Process (AHP).

INTRODUCTION

The trend of 4.0 is now beyond "Manufacturing" but could be seen everywhere including in the Tourism industry. Hotel 4.0 presents the future generation of the hotels while the hotels could be seen as an Intelligent Cyber-physical System. To make this dream happens, there are many aspects but technology is the core pillar. The transition process to shift to Hotel 4.0 is not always successful and the potential challenges are important to be recognized and considered in advance. The relation between ICT and the hotel industry has a long history but it's more important these days. Evaluating Information Technology

(IT) in United State hotels considering different hotel sectors from luxurious to economy ones has been studied by [1]. [2, 3] reviewed the existing research works on the implementation of quality programs in hotel industries located in Ireland. The relationship between hotel front office systems and the hotel guests has been studied by [4] considering the quality, information technology infrastructure, technology acceptance by guests and perceived value. [5] considered the implementation of customer relationship management (CRM) in 128 medium and small size hotels in Spain. The impact of hotel property attributes on adoption of RFID for hotel sector has been investigated by [6]. The potential barriers of implementing environmental technologies in the hotel industry in Hong Kong has been addressed by [7] while the paper categorized barriers into internal, external and product-related groups, and addressed some potential solutions. Reference [8] addressed managerial and theoretical extension of technology acceptance model based on evaluating the additional antecedent beliefs and predicted the attitude of tourists' behavior facing self-service technologies in the hotel environment. The key factors of CRM based on both customer-centered management, employee support and organizational culture have been addressed in [9]. [10] focused on the implementation of RFID used in hotel's locking technology to understand the adoption process and its theoretical implication. Possible barriers of applying information technology in hotels with the potential strategies as a solution have been addressed in [11]. Investigating existing related research works shows that the literature suffers lack of study addressing the potential practical challenges in implementing innovative technology. . This paper is going to point out these challenges and ranks them according to real case experiences. In this paper, the experience of Optishower, to Implement Innovative Water and Electricity Monitoring Technologies in hotels, is used to extract, categorize and rank the main challenges. The result is a combination of different Decision Process techniques including brainstorming, Delphi and Analytical Hierarchy process (AHP).

Optishower is a tech-based solution, offering an innovative platform for the hospitality. Optishower is a combination of hardware and software while the consumption of water and electricity is presented to hotel manager/associates via web dashboard. Although Optishower was successfully implemented in different European countries, still the probability of project failure is high. That's why dealing with challenges parallel with technological issues is part of the concern of the team. It is important to have a comprehensive picture of potential barriers ahead and evaluate the readiness of hotels in advance to develop an appropriate implementation strategy.

In this paper, the main barriers from previous project experiences were extracted through interviewing different groups involved in the project while Delphi was used to finalize the list. Afterwards they were ranked by using AHP technique to understand their priorities and importance.

CHALLENGES

To have a list of challenges, both service providers and client sides should be considered. Some challenges are based on the context and the place of implementation which could be different from one case to the other while the main dilemmas common among different case studies are the target of his work. Needless to say, that some difficulties were from the service provider side including technical problems and calibration among others which mainly deal with the technology itself. It should be noted that because of page limitation, the challenges are explained briefly.

Service Provider Challenges

Water Consumption Measurement

Water consumption can be measured in different I) intrusive and II) non-intrusive ways. Non-intrusive is more expensive and more difficult to code.

Intrusive Installation: In our experience, hall-effect water flow sensor was used because they were cheaper, easy to code and their accuracy was acceptable for the measurement. However, because of their intrusive technology for installation, we faced some problems such as lack of space for installation or difficulty to access to the water pipeline inside the rooms and the need to cut-off the water pipeline on the whole floor which is risky, especially when the hotel rooms are fully booked.

Hotel Renovation: Because of renovation in some hotels, even the hotel technicians do not know which pipe goes to which room, and it takes a lot of time to figure out how the system works.

Old and Worn-out Water Pipelines: Most of the water pipes and main valves were open for years and require significant efforts to close them for maintenance and installation. In addition, flow meter installation needs to be done by certified plumbers who are paid on an hourly basis and the complex pipelines increases the installation costs, while in some cases plumbers do not take the responsibility of the process.

Electricity Consumption Measurement

Insufficient space in room electricity box: The electricity box in many cases was too small and there was not enough space for installation.

Hotel Renovation: Similar to the challenges of water measurement, lack of knowledge about new electricity wiring system in renovated hotels is considerable.

Data Transfer Technology

There are different existing data transfer technologies such as Wi-Fi, Z-Wave and Zigbee. Among those mentioned technologies, Wi-Fi data transfer through Node MCU modules with ESP8266 Wi-Fi chipset is very popular, inexpensive, with lots of available libraries and source codes on the web. In addition, Wi-Fi infrastructure are always available in any hospitality areas. However, there are some challenges in this regard as follows:

Need for Power Supply: In case of providing power supply for water consumption metering and transmission, in most cases there was no accessible power line near the place where water flow sensors were needed to be installed.

Hotel's Wi-Fi Infrastructure: When you use Wi-Fi as a data transfer technology at least you need one router nearby, to be used and to connect your Wi-Fi modules to the cloud. Most of the hotels, according to our experience, use captive login to provide free Wi-Fi for their guests and it is hard to ask them for SSID and Password.

Power Outage: Having a power outage in long time monitoring is inevitable. In case of having short power outage, we implement a simple auto-reset circuit to refresh our Wi-Fi modules and make them reconnect to the internet.

No Internet Situation: Without the Internet, data cannot be sent to the cloud-based server and alternative solutions should be utilized.

Marketing

Hotels are working with few trusted suppliers in order to reduce diversion in after sale services mean that it's easier for new players to deal with existing hotel suppliers.

Client challenges

Infrastructure: Energy and water projects strongly depend on infrastructure and many challenges including various types of water counters and pipe sizes and the lack of map for **the** electricity grid are related to the infrastructure.

Security: Data security for any IoT project has always been a challenge and the risk of data disclosure is irreparable.

Safety: Implementing new solutions/products in hotels always jeopardizes hotel safety, and usually hotel managers should accept all the risks. One important challenge is to convince hotel managers that risk is under control.

System adaptation: Integration and synchronization of new innovative technology with existing systems which are currently working in the hotel is always a headache.

Decision process: It is difficult to find the right or final decision maker in hotel. This means, the decision process for a new product or technological solution in hotels is complicated and requires approval from different departments.

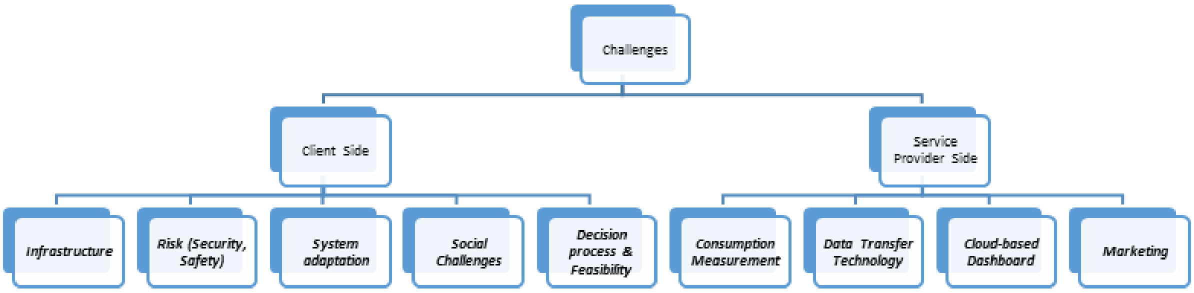
End user intervention: Hotel clients may intervene in any new product/solution while the hotel owner has no control over them. Vandalism or stealing stuffs in hotel rooms are two examples.

Economic Feasibility: The economic impact of a new technology/solution is not clear to the hotel managers and there is a need for a feasibility study or pilot duration to evaluate the product/Service which could cause delay or cost.

Technology Resistance: Resistance against new technologies is common and in the hotel industry, due to the sensitivity of the guests, it is worse.

RANKING AND WEIGTHING THE CHALLENGES

In this section, the process of ranking and weighting of the challenges using a well-known method, Analytical Hierarchy Process (AHP), is presented. Although all potential problems must be considered in any implementation strategy, the level of importance is not equal. After the initial steps by Delphi, to have the comparison tables of challenges, a brainstorming meeting was organized with the involvement of experts working in several projects. The first step was to develop a conceptual model in which we were trying to merge some criteria to reduce the complexity of the system and accordingly the possible errors in the weighing process. In AHP, the number of criteria is important and increasing the number of criteria, will dramatically affect the complexity and the risk of inconsistency. To avoid this problem, some similar challenges were merged. For instance, both water and electricity consumption measurement challenges could be merged to represent the consumption measurement challenge. In fig.1 the final model which represents the challenges could be seen.



Conceptual Model for AHP

To start the process of AHP, Pairwise Comparison scale adopted from Saaty [12] is used. Tables 1 and 2 represent the pairwise comparison matrix of criteria for challenges on both service provider & client sides respectively. These two tables are the final results of the brainstorming session. Although the opinions were not the same, ultimately, consensus was made to have the final tables.

Table 1. Pairwise comparison matrix of criteria for Service Provider Side Challenges

Service Provider Side Challenges	*Consumption Measurement*	*Data Transfer Technology*	*Cloud-based Dashboard*	*Marketing*
Consumption Measurement	1	6	2	5
Data Transfer Technology	0.143	1	0.2	5
Cloud-based Dashboard	0.5	5	1	7
Marketing	0.2	0.2	0.143	1

Table 2. Pairwise comparison matrix of criteria Client-Side challenges

Client-Side challenges	*Infrastructure*	*Risk (Security, Safety)*	*System adaptation*	*Social Challenges*	*Decision process & Feasibility*
Infrastructure	1	3	0.25	9	5
Risk (Security, Safety)	0.33	1	6	9	7
System adaptation	4	0.17	1	4	1
Social Challenges	0.11	0.11	0.25	1	0.2
Decision process & Feasibility	0.2	0.14	1	5	1

By following the process of "approximate method" in AHP (ref), the final result for the priorities of the criteria for the challenge are calculated and shown in table 3 and table 4.

Table 3. Priorities of Service Provider Side Challenges

Service Provider Side Challenges	**Consumption Measurement**	**Data Transfer Technology**	**Cloud-based Dashboard**	**Marketing**	**Priority**
Consumption Measurement	1	7	2	5	0.487
Data Transfer Technology	0.143	1	0.2	5	0.123
Cloud-based Dashboard	0.5	5	1	7	0.335
Marketing	0.2	0.2	0.143	1	0.056

Table 4. Priorities of Client-Side challenges

Client-Side challenges	***Infrastruc ture***	***Risk (Security, Safety)***	***System adaptation***	***Social Challenges***	***Social Challenges***	
Infrastructure	1	3	0.25	9	5	0.31
Risk (Security, Safety)	0.33	1	6	9	7	0.36
System adaptation	4	0.17	1	4	1	0.22
Social Challenges	0.11	0.11	0.25	1	0.2	0.03
Decision process & Feasibility	0.2	0.14	1	5	1	0.09

As it can be seen, the main challenge in Service Provider Side is consumption measurement following by cloud-based dashboard challenge while in client side, infrastructure and risk (security, safety) are considered as the main dilemmas followed by system adaptation.

One of the main concerns in AHP process is consistency. This is the way to make sure the judgments made by experts are consistent. The decision makers are using pairwise comparison and the criteria are compared with each other one by one, so some inconsistency is expected, and it is important to make sure the level of inconsistency is acceptable. To calculate the inconsistency, the main process which is introduced by Saaty [12] was followed as follows. λ_{max} for both Service Provider Side and client-side challenge are calculated by the process which could be seen in Table 5 and Table 6.

Table 5. Priorities as factors and weighed columns of Service Provider Side Challenges

Service Provider Side Challenges	***Consumpti on Measurem***	***Data Transfer Technolog***	***Cloud-based Dashboard***	***Marketing***	***Weighted sum***	***Priority***	***Weighted sum/ Priority***
Consumption Measurement	0.49	0.86	0.67	0.28	2.30	0.49	4.71

Data Transfer Technology	0.07	0.12	0.07	0.28	0.54	0.12	4.37
Cloud-based Dashboard	0.24	0.61	0.33	0.39	1.58	0.33	4.72
Marketing	0.10	0.02	0.05	0.06	0.23	0.06	4.06
Total	17.86	***Divide Total by 4 to obtain*** $\lambda_{\max}$					4.47

Table 6. Priorities as factors and weighed columns of Client-Side challenges

Client-Side challenges	***Infrastructure***	***Risk (Security, Safety)***	***System adaptation***	***Social Challenges***	***Decision process & Feasibility***	***Weighted sum***	***Priority***	***Weighted sum/ Priority***
Infrastructure	0.31	1.08	0.05	0.22	0.43	2.11	0.31	6.76
Risk (Security, Safety)	0.10	0.36	1.29	0.22	0.61	2.59	0.36	7.17
System adaptation	1.25	0.06	0.22	0.10	0.09	4.70	0.22	21.80
Social Challenges	0.03	0.04	0.05	0.02	0.02	0.17	0.02	6.89
Decision process & Feasibility	0.06	0.05	0.22	0.12	0.09	0.54	0.09	6.22
Total								48.82
Divide Total by 5 to obtain $\lambda_{\max}$								9.76

Now the Consistency Index can be calculated using (1) when N is 4 for service provider side and five for client-side challenge.

$$CI = (\lambda_{\max} - n) / (n - 1) \quad (1)$$

And finally, the Consistency ratio will be calculated using (2) while RI is the consistency index of a randomly generated comparison matrix. The RI index is the average of consistency Index of 500 randomly filled-in matrices [12]. RI index according to the number of criteria for N=4 is 0.9 and for N=5 is 1.12 [13, 14].

$$CR = CI / RI \quad (2)$$

Table 7. Consistency result for Service provider and client sides challenge

fa	$\lambda_{\max}$	***n***	***CI***	***RI***	***RI***
Service Provider Side Challenges	***4.47***	***4***	***0.16***	***0.9***	***0.17***
Client-Side challenges	***9.76***	***5***	***1.19***	***1.12***	***1.06***

Unfortunately, CR is bigger than 0.10. For values smaller than 0.10, it could be assumed that Judgment matrix is reasonably consistent. This means that the expert matrix for pairwise comparison was not acceptable and in order to have accurate data, it should be revised.

DISCUSSION AND CONCLUSION

This paper targeted an important subject of Technology Implementation challenges while trying to share the experiences from real projects. Industry 4.0 is a new paradigm and it's more than hard and soft technology but prosperous implementation strategy. Although the focus of this paper was Innovative Water and Electricity Monitoring Technologies in Hotel Industry, many factors could be the same in other industries and for other types of technology. Through interviews with the technical and commercial people involved in the projects, different challenges were identified and introduced while it was finalized by using Delphi Technique. Then AHP process was used to understand the weight of challenges. Although the final result was not satisfactory, due to the level of inconsistency, and the final weights were not accurate, but the ranking was interesting. Most probably the variety of the background of the contributors was the reason for deviation and low level of inconsistency. In Service provider side the main challenge is measurement system which is a technological challenge followed by cloud base dashboard. The result shows that marketing is not a big issue to deal with as compared with technology. In client side, the risk is the most important concern, which makes sense in hotel industry. They are sensitive and any risk in the safety or security could cause an irrecoverable cost. Infrastructure ranks the second and the existing situation of infrastructure and their potential problems even in 4-/5-star hotels were not expected. System adaptation stands third and most hotels suggest negotiations with a third party who is in charge of running and maintaining existing ICT systems in the Hotel. Also, communication between technical staff responsible for infrastructure and the ICT department caused some problems. They had different mindsets and different approaches. Social challenges in fourth place were more a concern of luxury hotels. This means that the gamification and the rewarding strategy does not work properly unless the hotel uses the technology align with their sustainability approach and green strategy. Contrary to our expectation, feasibility and decision process issues were in the last place. This clearly shows that the main challenges are not economic but either related to the technology or adaptation.
The project is still going on and the future works' concentration is on improving the current result by using other weighing techniques and introducing "I4.0 Readiness model" at micro level to understand the readiness of the hotels that would like to start transition strategy to be known as Tourism 4.0 Leaders. It is crucial to have a better understanding of the potential future challenges and plan appropriate customized implementation strategy according to the situation of each hotel and it's ecosystem.

ACKNOWLEDGMENT

This paper was partially financed by Portugal 2020, Madeira 14-20, Instituto de Desenvolvimento Empresarial, Região Autónoma da Madeira, for grants M1420-01-0247-FEDER-000004, with acronym: OPTISHOWER.

REFERENCES

[1] J. A. Siguaw, C. A. Enz and K. Namasivayam, "Adoption of Information Technology in US Hotels: Strategically Driven Objectives," Journal of travel Research, vol. 39, no. 2, pp. 192-201, 2000.

[2] M. Keating and D. Harrington, "The Challenges of Implementing Quality in the Irish Hotel Industry: A Review," Managing Service Quality: An International Journal, vol. 12, no. 5, pp. 303-315, 2002.

[3] M. Keating and D. Harrington, "The Challenges of Implementing Quality in the Irish Hotel Industry," Journal of European Industrial Training, vol. 27, no. 9, pp. 441-453, 2003.

[4] T. G. Kim, J. H. Lee and R. Law, "An Empirical Examination of the Acceptance Behaviour of Hotel Front Office Systems: An Extended Technology Acceptance Model," Tourism Management, vol. 29, no. 3, pp. 500-513, 2008.

[5] A. Padilla-Meléndez and A. Garrido-Moreno, "Customer Relationship Management in Hotels: Examining Critical Success Factors," Current Issues in Tourism, vol. 17, no. 5, pp. 387-396, 2014.

[6] A. Bulent Ozturk and M. Hancer, "Hotel and IT Decision-maker Characteristics and Information Technology Adoption Relationship in the Hotel Industry," Journal of Hospitality and Tourism Technology, vol. 5, no. 2, pp. 194-206, 2014.

[7] E. S. Chan, F. Okumus and W. Chan, "Barriers to Environmental Technology Adoption in Hotels," Journal of Hospitality & Tourism Research, vol. 8, no. 3, p. 1–25, 2015.

[8] A. K. Kaushik, A. K. Agrawal and Z. Rahman, "Tourist Behaviour Towards Self-service Hotel Technology Adoption: Trust and Subjective Norm as Key Antecedents," Tourism Management Perspectives, vol. 16, no. 1, pp. 278-289, 2015.

[9] N. Bahri-Ammari and K. Nusair, "Key Factors for a Successful Implementation of a Customer Relationship Management Technology in the Tunisian Hotel Sector," Journal of Hospitality and Tourism Technology, vol. 6, no. 3, pp. 271-287, 2015.

[10] L. M. Cobos, C. Mejia, A. B. Ozturk and Y. Wang, "A Technology Adoption and Implementation Process in an Independent Hotel Chain," International Journal of Hospitality Management, vol. 57, no. 1, pp. 93-105, 2016.

[11] F. Okumus, A. Bilgihan, A. B. Ozturk and X. Zhao, "Identifying and Overcoming Barriers to Deployment of Information Technology Projects in

Hotels," Journal of Organizational Change Management, vol. 30, no. 5, pp. 744-766, 2017.

[12] T. L. Saaty, Decision Making for Leaders: The Analytic Hierarchy Process for Decisions in a Complex World, Third Revised Edition. Pittsburgh: RWS Publications., 2012.

[13] M. Brunelli, Introduction to the Analytic Hierarchy Process, Springer, 2014.

Group DSS, Collaborative & Interactive Models

How to support group decision making in horticulture: An approach based on the combination of a centralized mathematical model and a Group Decision Support System

Pascale Zaraté, Guy Camilleri
IRIT, Toulouse University, Toulouse, France
{zarate,camiller}@irit.fr

MME Alemany, Ana Esteso Alvarez,
CIGIP, Universitat Politècnica de València, Camino de Vera S/N, 46002 Valencia, Spain
mareva@omp.upv.es, aesteso@cigip.upv.es

Mariana del Pino
FCAyF UNLP, La Anunciación - Huerta Orgánica, La Plata, Argentina
mdelpino@agro.unlp.edu.ar

ABSTRACT

Decision making for farms is a complex task. Farmers have to fix the price of their production but several parameters have to be taken into account: harvesting, seeds, ground, season etc… This task is even more difficult when a group of farmers must make the decision. Generally, optimization models support the farmers to find no dominated solutions, but the problem remains difficult if they have to agree on one solution. In order to support the farmers for this complex decision we combine two approaches. We firstly generate a set of no dominated solutions thanks to a centralized optimization model. Based on this set of solution we then used a Group Decision Support System called GRUS for choosing the best solution for the group of farmers. The combined approach allows us to determine the best solution for the group in a consensual way. This combination of approaches is very innovative for the Agriculture domain.

Keywords: Centralized Optimization Model, Group Decision Support System, AgriBusiness.

Find and cite the full work as:
Zaraté, P., Alemany, M., Del Pino, M., Esteso Alavarez, A.E., Camilleri, G., 2019. How to support group decision making in horticulture: An approach based on the combination of a centralized mathematical model and a Group Decision Support System, in: Freitas, P.S.A., Dargam, F., Moreno, J.M. (Eds.), Decision Support Systems IX: DSS Developments & Future Trends, Lecture Notes in Business Information Processing. Springer International Publishing, Cham, pp. 78-89.

Collaborative value modelling in widely participated decision support processes

Carlos A. Bana e Costa, Ana C.L. Vieira, Mónica D. Oliveira
CEG-IST, Centre for Management Studies of Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal
carlosbana@tecnico.ulisboa.pt, monica.oliveira@tecnico.ulisboa.pt, ana.lopes.vieira@tecnico.ulisboa.pt
João Bana e Costa
Decision Eyes, Rua da Prata, 80, Lisboa, Portugal
joao.bana@decisionseyes.com

ABSTRACT

Departing from the classical collaborative knowledge acquisition methodology in expert systems literature, the 'collaborative value modelling' framework was developed. It combines two decision support processes: Web-Delphi to elicit individual judgemental knowledge from a (very) large number of stakeholders, and Decision Conferencing (DC) in which the knowledge acquired is then digested by a small group of key-players, to collaboratively develop a widely informed multicriteria value model. This new socio-technical framework was already tested in real complex evaluation contexts, supported by the Welphi platform and the MACBETH method. This paper focuses on the Welphi process for weighting population-health objectives that informed a subsequent multicriteria DC to construct a European Population Health Index under the EURO-HEALTHY H2020 project. It shows that enhancing MACBETH decision conferencing with an ex ante Web-Delphi process fostered higher participation and collaboration in multicriteria modelling. Furthermore, it shows how the Welphi platform, by collecting participants' changes of opinion (or judgements) between Delphi rounds, provides a mean to understand the knowledge construction process itself. Current research under the IMPACT HTA H2020 and MEDI-VALUE FCT projects addresses both technical and social challenges, on the one hand, to find ways for improving the platform in order to understand the reasons behind opinion changes and knowledge generated by participants' interactions and comments, and on the other hand, to explore behavioral aspects within and across Delphi, DC and MACBETH web-based group decision support systems.

Keywords: Collaborative modelling; Knowledge acquisition; Multicriteria decision conferencing; Participatory process; MACBETH; Web-Delphi; WELPHI

INTRODUCTION

Delphi and decision conferencing (DC) can be appropriately viewed as group decision support systems (DSS) [1], as they both make use of an information basis and decision support methods and technology to acquire individual knowledge and stimulate collaborative group agreement. From the social view point, the difference between these two particular DSS is that Delphi are non-face-to-face anonymous processes able to involve a large number of participants, whereas DC develops on open face-to-face settings restricted to relatively small groups. From the technical perspective, we will restrict the discussion of their use on decision aiding contexts which ultimate objective is the construction of multicriteria evaluation models under the working hypotheses of multi-attribute additive value modelling (MAVM) [2] and using the MACBETH approach [3, 4]. Specifically, we will mention the enchained use of the Welphi platform [5] and the M-MACBETH software [6] to develop, respectively, Web-Delphi and DC processes, framed on the recently proposed 'collaborative value modelling framework' [7].The next section is devoted to this novel socio-technical framework, followed by the description of its application to a real case of weighting population health objectives in Europe.

THE COLLABORATIVE VALUE MODELLING FRAMEWORK

The expert systems literature has recognized the value of including a large number of experts to help dealing with the increase complexity of current applications [8]. Collecting information from multiple sources (experts) has numerous benefits, such as an improved knowledge of the problem domain, and enhanced "knowledge base validity, consistency, completeness, accuracy, and relevancy," that will produce better results within multidimensional problems [9]. Wagner [10] in a recent review over the last Thirty Years of Expert System Case Studies identified different hybrid knowledge-based systems that combine multiple knowledge sources to develop expert systems that are currently exploring how different methodologies can be combined to face this challenge. However, the development of a collaborative knowledge acquisition process raises several challenges, particularly on how to use technological platforms to assist knowledge acquisition tasks that involve a large number of experts and stakeholders that can be geographically dispersed.

Drawing from the collaborative knowledge acquisition literature and from the idea that "it is possible to realize MAVM processes in which integration and interaction are at different levels" [11], we designed the above referred collaborative value modelling framework, a new socio-technical framework that departs from Web-Delphi processes to elicit and analyse individual judgemental knowledge from a (very) large and diverse number of stakeholders; the knowledge thus acquired is then digested by a small group of key-players, in a subsequent DC process, to collaboratively develop a widely informed multicriteria evaluation model.

A detailed description of the different stages of the Collaborative Value Modelling framework can be found in [7]. In this paper we focus on describing how the knowledge acquired flows between the two decision support processes used as environment components of the framework: Welphi and Decision conferencing (DC). Quoting Vieira et al. (2019) [7]: "At the heart [of the collaborative value modelling framework] are the multiple stakeholders whose involvement takes place in two different environments, assisted by the facilitation team: a participatory environment in the form of a non-face-to-face Web-Delphi, with a large number of participants; and a collaborative environment in the form of a face-to-face multicriteria decision conferencing, with a group of key-players."

Welphi was developed using the Microsoft ASP.Net framework supported by an SQL Server database. The front-end makes extensive use of JQuery and Ajax to provide an easy experience for the users. Welphi has three main sections of interaction. The first section is where the user sets up the process itself. This includes the definition of the indicators under study, the answer scale to be used throughout the process rounds and the addition of the participants e-mails. The second section lets the user setup, monitor and analyse each round. It is possible to edit invitation and reminder e-mails' bodies, start and stop the round as well access the participants answering status and the round statistics. Welphi also integrates the novel online MACBETH-based decision support system that also implements the MACBETH approach in a web setting environment [5].

A PROCESS OF COLLABORATIVE VALUE MODELLING FOR WEIGHTING

The EURO-HEALTHY case

The collaborative value modelling framework has been fully applied for developing a MAVM Population Health Index (PHI) model, under the EURO-HEALTHY H2020 project, to appraise PH across 269 European regions. Three separate Web-Delphi processes took place, always using the Welphi platform. Based on scientific information, the first one enabled to build a population-health value tree with 10 areas of concern, each one of them clustering multiple indicators. They operationalize a total of 42 population health objectives, which weighting process started with the second Web-Delphi. Finally, the shapes of the respective value-functions were the object of the third Web-Delphi. Given the multidimensional nature of the PH concept, many geographically dispersed experts and stakeholders from different areas were involved. Those three types of shared judgmental knowledge, thus collected, were analyzed and discussed by a group of 13 key-players, each type during a one-day decision conference to achieve agreement on, respectively, the PHI value-tree, the final weights of the objectives and the respective value functions. This paper focuses only on the weighting collaborative modelling activities, to illustrate how weighting judgements were friendly and comprehensively elicited from a large and diverse number of participants and how the weighting knowledge collected with Welphi informed the weighting decision conference.

Web-Delphi for weighting

The Web-Delphi for weighting was a one-month process that took place between February 29 and April 4, 2016. In three rounds, qualitative weighting judgements on the indicators that operationalize European health objectives were collected from 61 participants. The knowledge elicitation phase started with the design of the questionnaire of the first round of the Web-Delphi process (Figure 8). In weighting, it was important to consider that, technically, in an additive value model, each weight is a quantitative measure of the relative importance of swinging between lower and upper performance levels on the respective indicator. In the case, weighting judgements were elicited by asking the participants to answer to the question: “To reduce health inequalities in Europe, how important is to close this gap?” All gaps were shown on the Welphi platform and defined between the worst and the best performance levels on the respective indicator, across all European regions. In the beginning of the first Delphi round, participants accessed the list of indicators and the respective gaps and gave their individual qualitative weighting judgements by choosing one of the six MACBETH categories, from very weak to extreme, directly on the platform, while considering “how big is the gap and how important it is to develop policies to close the gap in view of reducing health inequalities in Europe” [12].

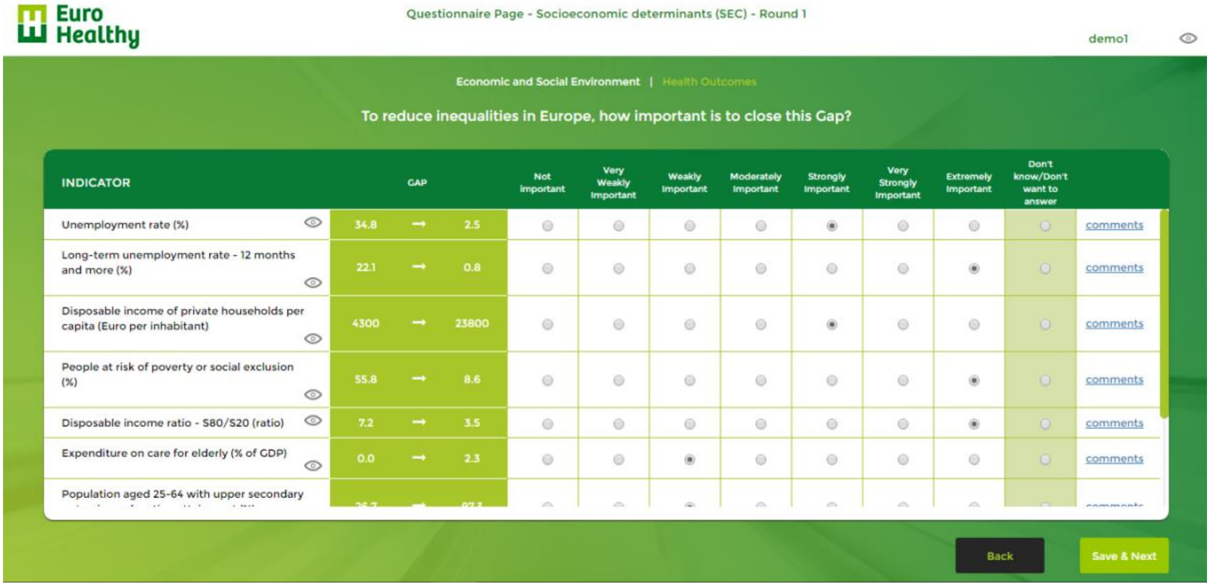


Figure 8: Web-Delphi for weighting in each area of concern: answers in the first round [12].

In the beginning of the second Delphi round, participants were presented with an anonymous summary of all participants' answers. It is this iterative nature of the Delphi method that enables the knowledge analysis phase of the collaborative value modelling framework. On the one hand, quantitative feedback was given in a table with each row indicating the percentage (and number) of participants' weighting judgements on each qualitative MACBETH category. It is important to mention that no statistical aggregation measure was delivered. This strategy, like a voting process as the one adopted in the [13] study, was more adequate to promote reflection. This was enhanced, on the other hand, by qualitative feedback in the form of comments that participants could give to justify their judgements. At the light of the feedback information provided, in round 2 participants were invited to confirm or revise their previous judgements. A similar process took place in a third and final round.

Individual opinion changes were used as measures of the usefulness of the Delphi method for collecting weights information. There were indicators where there was more individual opinion change on both rounds than in other indicators. The indicators with less opinion change, either started with agreement from the first round or no group majority was reached; whereas the indicators with more individual opinion changes reached a group majority judgement on round 2.

At the end, majority rules were applied to the final distributions of weighting judgements, in order to identify the indicators for which group majority judgements existed and those ones with controversy results.

Multicriteria decision conferencing for weighting

The one-day weighting decision conference took place, in Lisbon, on October 25, 2016, attended by 13 experts who had also participated in the preceding weighting Web-Delphi. Departing from the analysis and discussion of the qualitative weighting knowledge there acquired, the group developed the weighting model activities depicted in Figure 2 and moved forward the value modelling process by achieving an agreement on numerical weights to the indicators' gaps, with the support of M-MACBETH. When Delphi majority qualitative weighting judgements were not respected by the final quantitative weights agreed in the DC, clear justifications were asked in order to guarantee the transparence of the whole collaborative value modelling process developed. Within the DC process, participants later verified and validated the weights of the PHI model.

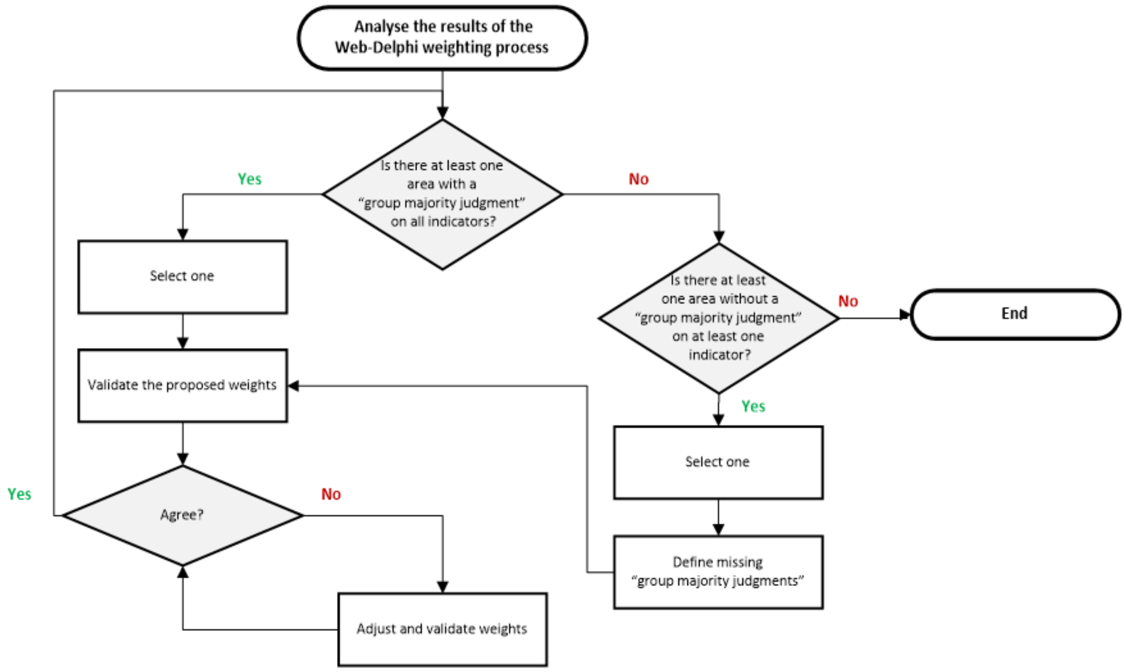


Figure 9: Flowchart of the weighting decision conference activities [14].

CONCLUSIONS

The Web-Delphi, as used in stage 2 of the collaborative value modelling framework, showed to be effective in "introducing knowledge acquisition as a cognitive process" [15]. This is in line with the state-of-the-art of knowledge acquisition. Furthermore, the way in which the Web-Delphi was designed allowed to capture disagreements among participants, while witnessing their evolution towards agreement – a feature usually only captured during face-to-face processes [16]). The process was also an efficient way to operationalise knowledge acquisition from 61 geographically dispersed participants, as it allowed for an automatic record of the information collected which saved facilitation time and avoided unnecessary errors, simultaneously making the information collected available for the subsequent stages of the process. Finally, having available previous knowledge information proved to increase the confidence of the DC group in developing their tasks while reducing the time generally allocated to them.

Analysis of data stability and individual opinion change are key aspects to understand participatory processes, particularly the Delphi method. Further technical research within the IMPACT HTA H2020 is now being conducted to explore the reasons behind opinion changes and the insights provided by Delphi participants' comments along the process. Further social research within the MEDI-VALUE FCT is being conducted to capture behavioural aspects that may occur within Delphi processes and that may be relevant to DC participants interpreting and understanding Delphi results. Both projects also address other socio-technical issues within and across Delphi, DC and MACBETH web-based group decision support systems.

REFERENCES

[1] DeSanctis, G. and R.B. Gallupe, *A Foundation for the Study of Group Decision Support Systems.* Management Science, 1987. **33**(5): p. 589-609.
[2] Winterfeldt, D.v. and W. Edwards, *Decision Analysis and Behavioral Research*. 1986: Cambridge University Press. 626.
[3] Bana e Costa, C.A., J.-M. De Corte, and J.-C. Vansnick, *Macbeth.* International Journal of Information Technology & Decision Making, 2012. **11**(02): p. 359-387.
[4] Bana e Costa, C.A. and J.-C. Vansnick, *MACBETH — An interactive path towards the construction of cardinal value functions.* International Transactions in Operational Research, 1994. **1**(4): p. 489-500.
[5] Welphi. *Welphi*. 2017; Available from: http://www.welphi.com/.
[6] M-MACBETH . *M-MACBETH*. 2017; Available from: http://m-macbeth.com/.
[7] Vieira, A.C.L., M.D. Oliveira, and C.A. Bana e Costa, *Enhancing knowledge construction processes within multicriteria decision analysis: The Collaborative Value Modelling framework.* Omega, 2019.
[8] Wagner, W.P., M.K. Najdawi, and Q.b. Chung, *Selection of knowledge acquisition techniques based upon the problem domain characteristics of production and operations management expert systems.* Expert Systems, 2001. **18**(2): p. 76-87.
[9] Medsker, L., M. Tan, and E. Turban, *Knowledge acquisition from multiple experts: Problems and issues.* Expert Systems with Applications, 1995. **9**(1): p. 35-40.
[10] Wagner, W.P., *Trends in expert system development: A longitudinal content analysis of over thirty years of expert system case studies.* Expert Systems with Applications, 2017. **76**: p. 85-96.
[11] Marttunen, M., et al., *How to design and realize participation of stakeholders in MCDA processes? A framework for selecting an appropriate approach.* EURO Journal on Decision Processes, 2015. **3**(1-2): p. 187-214.
[12] EURO-HEALTHY WP6, *Technical report on scenario building and analysis of policies*. 2017: Instituto Superior Técnico - Universidade de Lisboa. p. 197.
[13] Bana E Costa, C.A., et al., *A socio-technical approach for group decision support in public strategic planning: the Pernambuco PPA case.* Group Decision and Negotiation, 2014. **23**(1): p. 5-29.
[14] EURO-HEALTHY WP6, *Development of the multicriteria model to evaluate Population Health*. 2016: Instituto Superior Técnico - Universidade de Lisboa. p. 168.
[15] Cairó, O., *KAMET: A comprehensive methodology for knowledge acquisition from multiple knowledge sources.* Expert Systems with Applications, 1998. **14**(1): p. 1-16.
[16] Vivacqua, A.S., A.C.B. Garcia, and A. Gomes, *BOO: Behavior-oriented ontology to describe participant dynamics in collocated design meetings.* Expert Systems with Applications, 2011. **38**(2): p. 1139-1147.

A search-tree approach to generate project portfolios in the presence of interactions

Rudolf Vetschera
University of Vienna
Oskar Morgenstern Platz 1, 1090 Vienna, Austria
Rudolf.vetschera@univie.ac.at
https://bda.univie.ac.at

Jonatas Almeida
Universidade Federal de Pernambuco, CDSID - Center for Decision Systems and Information Development
Recife, Brazil
jonatasaa@yahoo.com.br
http://www.cdsid.org.br

ABSTRACT

We consider the problem of selecting a portfolio of projects which exhibit possible synergies in their resource consumption as well as outcomes. We propose a tree based search process to generate the set of efficient portfolios in a multi-criteria context and develop bounds on objective values that can be used to limit the search process in the tree. Initial computational results show that reasonable tight bounds can be generated using a general approach, for which significant calculations can be performed ex ante.

Keywords: Project portfolio; portfolio decision analysis; interactions; search tree; bounds; knapsack

INTRODUCTION

Decision analysis has for a long time dealt with problems that consider decision alternatives individually, such as choice problems in which a single alternative has to be chosen, or ranking and sorting problems, in which individual alternatives are ranked or assigned to predefined categories [1]. More recently, researchers have begun to study problems of portfolio decision analysis [2], in which several alternatives can be implemented simultaneously. Portfolio decision problems are relevant in many applications, a prominent example is management of projects such as R&D projects [3].

As long as only one alternative can be selected, interdependencies between alternatives are not relevant. However, this is no longer true in portfolio problems, where several projects can be realized together. Early models of portfolio decisions and project selection [4] were mainly concerned with interdependencies in the implementation of projects, e.g. that one project can only be carried out if another project is also carried out, or that two projects are mutually exclusive. In the present paper, we deal with interactions in the form of positive or negative synergies of projects. For example, execution of one project might have a positive or negative impact on the payoff of another project, or two projects might be able to share some resources, leading to synergies in their resource consumption.

Our aim is to develop models to support a decision maker in portfolio problems in which projects might exhibit positive or negative synergies both in terms of their resource utilization, and in their contributions to the decision maker's objectives. A portfolio decision problem can be seen as a combinatorial decision problem [3]. In the single objective case, it can be solved by a knapsack approach [5]. If interdependencies between projects result solely from the fact that they all require resources from a common pool, and there are no synergies in resource requirements or the utility of projects, this approach can directly be extended to multiattribute problems by assigning a utility value to each project [6]. However, this requires strong assumptions about the decision maker's preferences, as well as an a priori elicitation of the decision maker's multi-attribute utility function that may be imprecise and/or cognitively hard for him/her [7, 8]. A frequently used approach that avoids these problems consists of first generating the set of efficient portfolios, and then using an interactive search procedure in this (discrete, but possibly large) set to identify the best portfolio [9]. However, even in an offline environment, it is not possible to generate all possible portfolios to identify the efficient ones. Thus, tests for efficiency have to be included in the generation process to restrict the search space. This requires that during the search process, some information about outcomes that can be reached within a certain part of the search tree is available. The search space can be restricted even further by incorporating preference information from the decision maker in an interactive way, which also requires information about possible outcomes. Thus, the problem of calculating bounds on objectives during the search process becomes crucial. In this paper, we formulate such bounds and evaluate them in a computational study.

The remainder of the paper is structured as follows. In the second section, we present our model of interactions between projects in portfolios. Section three outlines a tree-based search process for efficient portfolios and develops formulas for bounds on objectives. Section four presents results from a computational study on the efficiency of these bounds and section five concludes the paper.

INTERACTIONS

We propose to represent interactions between projects in a way that is similar to interactions in the capacity function of a Choquet integral [10]. The same approach can be used to model interactions in the objective functions as well as constraints (resource requirements). Let S be a set of projects, i.e., a portfolio. This can be the entire set of projects available to the decision maker, or a subset thereof. Both the (multiple) objective functions and the resource requirements need to be evaluated at the level of entire portfolios. We define a function $f(S)$ that considers interactions at the portfolio level as

$$f(S) = \sum_{i \in S} v_i + \frac{1}{2} \sum_{i \in S} \sum_{j \in S} c_{ij} \quad (1)$$

where v_i refers to the outcome (objective value or resource requirement) of project i if the project is carried out alone and c_{ij} represents a positive or negative synergy effect, i.e., the amount by which the outcome of the portfolio deviates from the sum of outcomes of the two projects if both projects i and j are contained in the portfolio. Note that the interaction effect has to be divided by two to avoid double counting.

We call a function f *monotonic in projects* if adding a new project to a portfolio does not decrease f. Formally, function f is monotonic if

$$\forall i \notin S : f(S \cup \{i\}) \geq f(S) \quad (2)$$

Although it is hard to imagine that interactions are so strong that adding another project to a portfolio would decrease its outcome or the amount of some resource needed, we do not explicitly exclude the possibility that a function is not monotonic.

For the moment, we consider only a single objective. The problem of selecting the optimal portfolio can then be formulated as a knapsack problem:

$$\max \sum_{i\in S} v_i^0 x_i + \frac{1}{2}\sum_{i\in S}\sum_{j\in S} c_{ij}^0 x_i x_j$$
$$\sum_{i\in S} v_i^k x_i + \frac{1}{2}\sum_{i\in S}\sum_{j\in S} c_{ij}^k x_i x_j \leq b^k \qquad (3)$$

where x_i is a binary variable indicating that project i is included in the portfolio and b^k is the available amount of resource k. Note that although (3) is a nonlinear optimization model, it can be transformed into a linear model by substituting $x_i x_j = y_{ij}$ and adding the constraints

$$y_{ij} \leq x_i$$
$$y_{ij} \leq x_j$$
$$y_{ij} \geq x_i + x_j - 1 \qquad (4)$$

Similar formulations of optimization models taking into account project interdependencies were presented e.g. in [11] and [12], although they did not consider the linearization.

SEARCH TREE AND BOUNDS

For the efficient generation of portfolios, we propose to build up a search tree similar to [13]. The tree is constructed by considering each project sequentially. For each considered project, two branches are added to the tree, one in which the project is included in the portfolio under construction and one in which it is not included. The end nodes of the tree correspond to complete portfolios.

The advantage of a tree structure is that it is not necessary to search the entire tree. If at some point it can be determined that all nodes descending from a node will violate a resource constraint, or that none of the following nodes will be better than a solution already found (or, in the multiobjective case, all subsequent nodes will be dominated by an existing solution), then the tree need not be further expanded from that node. The efficiency of this approach depends on the sequence in which projects are considered, different approaches to select projects will be studied in future research.

In the case that all resource requirements are monotonic, it is obvious that once a resource requirement has been violated, all subsequent nodes will also violate that requirement. This no longer holds in the case of non-monotonic resource requirements. In that case, a bound on resource usage is required. It is also necessary to calculate a bound on objective values in order to determine whether nodes that emerge from a given node can still be better than a solution found so far. Thus, the ability to calculate tight bounds on both resource usage and objective values of all nodes emerging from a given node becomes essential

The tree search uses a trichotomic approach as proposed in [14]. Consider a given node n in the tree, and denote by S_1 the set of all projects that are included in the portfolio under construction at node n, by S_2 the set of all projects that are already decided and that are not included in the portfolio, and by S_3 the set of still undecided projects. Although we can calculate the value of the objectives (as well as the resource requirements) for all projects in S_1 using (1), that value might still increase or (in the case of a non-monotonic function) decrease by adding projects from set S_3. We also have to take into account that adding too many projects might lead to a violation of the resource constraints.

To obtain an exact bound for an objective at node n, one could still solve problem (3) fixing x_i to one for all $i \in S_1$ and to zero for all $i \in S_2$. However, that would involve solving several mixed integer optimization problems (one for each objective) in each node. To efficiently calculate a bound for an objective function at node n, we therefore need to relax some properties of the problem to allow for fast calculation. We propose to ignore the indivisibility of projects, and to consider each resource constraint individually. Without interactions, a bound could then be calculated by considering, for each resource k, the marginal productivity of each project i

(which we denote by $r_{ik} = v_i^0 / v_i^k$), sorting the projects by productivity and including (partial) projects in descending order of productivity until the resource constraint becomes binding. Taking the smallest value that can be obtained in this way for any resource gives an upper bound for the objective value that can be reached.
If there are interactions between projects, the marginal productivity of a project depends on the other projects contained in the portfolio. Furthermore, in the case of non-monotonic functions, adding a project might lead to a decrease in either resource requirements or objective value, so the marginal effect might be negative. The latter effect can be eliminated by considering only positive interactions for objectives, and only negative interactions for resources. Thus, the marginal productivity of a project will only be over-estimated, leading to an (less tight) upper bound of the objective value that can be obtained.
At a given node, it is not clear which interactions of a project need to be considered. Clearly, all interactions of the project with other projects contained in S_1 will be present, and interactions with projects in S_2 will not be present. But interactions with other projects in S_3 may or may not occur. To be on the safe side, we only include positive interactions in the objective and negative interactions in resource usage with all projects still in S_3. Therefore, an upper bound of the marginal productivity of an undecided project $i \in S_3$ is calculated as

$$r_{ik} = \left(v_i^0 + \sum_{j \in S_1} c_{ij}^0 + \sum_{j: j \in S_3 \wedge c_{ij}^0 > 0} c_{ij}^0\right) \Big/ \left(v_i^k + \sum_{j \in S_1} c_{ij}^k + \sum_{j: j \in S_3 \wedge c_{ij}^k < 0} c_{ij}^k\right) \qquad (5)$$

Note that this bound depends on the current sets S_1 and S_3. This means that in each node, the marginal productivity of each project in S_3 has to be calculated, and the marginal productivities need to be sorted in each node. While faster than solving the entire problem (3), these operations are still quite time consuming. We therefore also consider an even more simplified approximation of the marginal productivity of a project, which utilizes all positive interactions in the objective and all negative interactions in the resources:

$$s_{ik} = \left(v_i^0 + \sum_{j \neq i \wedge c_{ij}^0 > 0} c_{ij}^0\right) \Big/ \left(v_i^k + \sum_{j \neq i \wedge c_{ij}^k < 0} c_{ij}^k\right) \qquad (6)$$

We refer to the bound calculated using r_{ik} as the *specific* bound at node *n*, and the bound calculated using s_{ik} as the *general* bound. Obviously, the specific bound will be tighter than the general bound.

COMPUTATIONAL RESULTS

In order to compare the two bounds defined in the previous section to each other, and to the actual best performance that could be obtained by solving model (3) for each node in the search tree, we conducted a computational study. All experiments involved problems with 10 projects and 3 resources. In the base case, the interaction matrix contained 50% of non-zero interactions between -10% and +10% of the maximum value. 300 randomly generated problems were analyzed. For each problem, we generated a search tree by considering the projects in the order of their sequence numbers. For each node in the tree, both the specific and the general bound of the objective were calculated, as well as the optimal value resulting from model (3) when fixing the variables of already decided projects.
Figure 1 shows the ratio of average bounds to the average optimal solution for different levels of the tree (i.e., different numbers of still undecided projects).

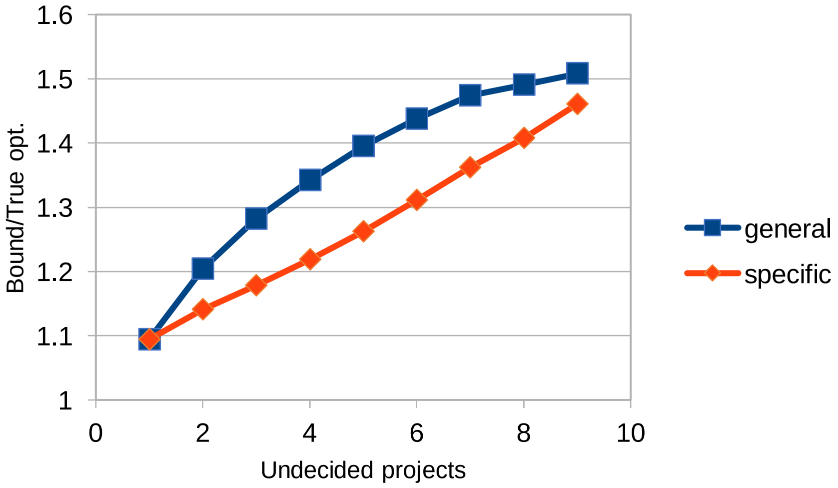


Figure 1: Ratio of bound to true optimum for different numbers of undecided projects

As could be expected, the bounds are tighter when only few projects are left undecided. But even when only one project is already decided, the bounds on average stay within 50% of the optimal value. Interestingly, the general bounds, which require considerably less computational effort, lead to an error which is at most 15 percentage points higher than the specific bounds.
To test the sensitivity of these results to problem parameters, we also performed simulations with different densities of the interaction matrix (20% and 90%), and stronger interaction effects (up to 50% of total value), and tighter or less tight resource constraints. As could be expected, the bounds become less tight the more and the larger the interactions are, but are more tight for less constrained problems. This is plausible since the set of undecided projects (S3) of the early nodes is more similar to the set of accepted projects of the efficient portfolios when the constraints are less tight.

CONCLUSIONS

In this paper, we discussed a tree-based search process for project portfolio selection problems when projects may exhibit synergies in terms of resource usage and outcomes. This process can be used for an interactive exploration of the set of efficient portfolios, for example, decision makers can set aspiration levels for objectives to limit the part of the tree that is being searched. To support this approach, it is necessary to obtain reliable bounds on objective values for partially constructed portfolios. We have shown that such bounds can be obtained with reasonable effort and provide good approximations of possible objective values. This makes it possible to conduct an efficient generation and search process. More detailed simulations on various aspects of this approach will be performed in future studies.

REFERENCES

[1] Roy, B., "Problems and methods with multiple objectives". Mathematical Programming, 1971. **1**: p. 239-266.
[2] Salo, A., J. Keisler, and A. Morton, eds. *Portfolio decision analysis*. 2011, Springer: New York.
[3] Heidenberger, K. and C. Stummer, "Research and development project selection and resource allocation: A review of quantitative modelling approaches". International Journal of Management Reviews, 1999. **1**(2): p. 197-224.
[4] Stummer, C. and K. Heidenberger, "Interactive R&D portfolio analysis with project interdependencies and time profiles of multiple objectives". IEEE Transactions on Engineering Management, 2003. **50**(2): p. 175-183.
[5] Eilon, S., "Application of the knapsack model for budgeting". OMEGA Int. J. of Management Science, 1987. **15**(6): p. 489-494.
[6] Kleinmuntz, D.N., *Resource allocation decisions*, in *Advances in decision analysis*, W. Edwards, R.F. Miles, and D. von Winterfeldt, Editors. 2007, Cambridge University Press: New York. p. 400-418.
[7] Keeney, R.L. and H. Raiffa, "Decisions with multiple objectives: Preferences and value tradeoffs". 1976, J. Wiley & Sons, New York.
[8] Vilkkumaa, E., J. Liesiö, and A. Salo, "Optimal strategies for selecting project portfolios using uncertain value estimates". European Journal of Operational Research, 2014. **233**(3): p. 772-783.
[9] Gettinger, J., et al., "A comparison of representations for discrete multi-criteria decision problems". Decision Support Systems, 2013. **54**: p. 976–985.
[10] Grabisch, M., "The application of fuzzy integrals in multicriteria decision making". European Journal of Operational Research, 1996. **89**(3): p. 445-456.
[11] Klapka, J. and P. Pinos, "Decision support system for multicriterial R&D and information systems projects selection". European Journal of Operational Research, 2002. **140**(2): p. 434-446.
[12] de Almeida, A.T. and M.D.O. Duarte, "A multi-criteria decision model for selecting project portfolio with consideration being given to a new concept for synergies". Pesquisa Operacional, 2011. **31**(2): p. 301-318.
[13] Vetschera, R., "Composite alternatives in group decision support". Annals of Operations Research, 1994. **51**: p. 197-215.
[14] Mavrotas , G. and O. Pechak, "The trichotomic approach for dealing with uncertainty in project portfolio selection: Combining MCDA, mathematical programming and Monte Carlo simulation". International Journal of Multicriteria Decision Making, 2013. **3**(1): p. 79-96.

Multi-Agents & Data Mining Approaches

A Multi Agent Architecture for IT Infrastructure Diagnostic

Elandaloussi Sidahmed
LIO / Ahmed Ben Bella Oran1 University
BP 1524 M'Naouer, Oran 31000, Algeria,
IRIT / Toulouse University
2 Rue du Doyen Gabriel Marty, 31042 Toulouse Cedex 9, France
elan.sidahmed@gmail.com, sid.elandaloussi@irit.fr
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
Taghezout Noria
LIO / Ahmed Ben Bella Oran1 University
BP 1524 M'Naouer, Oran 31000, Algeria
Taghezout.nora@gmail.com

ABSTRACT

IT infrastructure refers to all the underlying elements that necessary provide technological services to an end-user. Besides, it allows an organization which delivers IT solutions and services to its employees, partners and/or customers. In an IT infrastructure, the interruption service can result in significant financial penalties and/or a loss of users trust. So to ensure that our IT equipment (hardware, software, network…etc) is available to perform the expected level, we need to be proactive in detecting incidents that disrupt the proper functioning of the company. The main objective of our work is to propose an efficient diagnostic platform for an IT infrastructure called (**MAITD)** in order to assist the company technicians to have a pertinent solution(s) in a short time face to decision-making problem.

Keywords: Decision Support, Multi-Agent System, IT Infrastructure, Corrective Maintenance.

INTRODUCTION

IT infrastructure maintenance is the key that ensures business continuity which can be divided into several areas: server, storage, network, security, virtualization…etc. There may be some specific maintenance requirements for particular IT(Information technology) infrastructures that should be defined by the organization itself. In addition, most of the maintenance operations are conducted through preventive methods with fixed time [1] or corrective [2]. In recent years the purpose of the maintenance task is to calculate the maintenance needs before the equipment failure, i.e. a continuous monitoring [3]. Several proprietary software solutions around IT infrastructure environment have appeared in this context such as: Solar Winds[4], Manage Engine[5], Zenoss[6]. However, in overall are preventive solutions and can't produce solution if infrastructure is affected. Additionally, decision supports systems have been developed to assist decision makers to solve their problems in a preventive or corrective manner. Our study is a part of a corrective approach in order to find a better solution to a problem in a short time. Moreover, the description problems have never been considered perfect according to the user expertise level and also around the

database structure that will be interrogated (keywords, structured query language, dictionary…etc). Nevertheless, this fact leads us to find unsatisfactory solutions, as consequence, we must treat the entered data by the user differently before starting the solution process.
This paper is organized as follows: Section two presents the different strategies for maintenance with some related works. Section three describes the proposed approach by introducing our general architecture. Section four shows some implementation examples of the case study during troubleshooting task. Finally, section 5 summarizes the paper and offers directions for future work.

RELATED WORK

There have been many works related to the information technology maintenance operations such as: the works presented in [7], which propose the taxonomy to outline modern intrusion detection systems IDSs. In addition, the work in [8] presented a Web and Agent Technologies in monitoring and maintenance condition of mechanical and electrical systems.
The multi-Agents systems have been used successfully in many application areas as with maintenance field. Indeed, several searches have been realized for example: the work described in [9], which used a hierarchical Multi-Agent System (MAS) for monitoring and reporting policy violations within the security environment. [10] proposed a decentralized multi-agent security system (DMASS) as a scalable solution for the collection and analysis of cyber security and network forensic data. Authors in [11] introduced an agent framework for forensic information gathering by using three types of agents for data collection, analysis and alert generation.
All these researches are preventive in a special way to take a fundamental step to avoid any breakdowns.
On the other hand, there are some surveys dealt in corrective maintenance but only for a security point of view like Work in [12], proposed algorithm considering multiple attributes of user keystroke dynamics, which in addition of traditional authentication has applied in an organization for distinguishing one user than another.
Our approach is a corrective one based on the support basically focus on corrective information technology infrastructure maintenance as with [13], which proposed a decision support framework in an IT environment which was essentially based on agents' mobiles and web services. However, the main problem with mobile agent is security which is still an area of research on its own. Indeed, the latter work is limited to ensure a secure access management. In this paper, a multi agent system is proposed as able to coordinate and cooperate his action in order to have a corrective decision method for IT infrastructure maintenance as it is shows on the next section.

PROPOSED APPROACH

Our **MAITD** architecture is a multi-layered architecture namely: Presentation, Interpretation, and Data Layer. The highest layer is constituted of two agents' categories, specifically the participant and expert agent who act as interface with the system in order to treat all input or output data. Additionally, the middle layer contains the rest of agents which is considered as kernel in our architecture with their functions and methods that are implemented in java environment. Finally, the lowest layer groups together all data source which are requested. In the following, we described main tasks of each agent:
Participant Agent: Treats the input information of the user through breaking down the problems in pertinent terms (Bags of Words) by using a Tree Tagger tool and our knowledge

bases. Also, it is responsible for forwarding all system notifications to the corresponding participant.

Expert Agent: Represents a corresponding expert with his agenda, preference, and profiling information. So, it is able to accept or refuse any meeting invitation. Furthermore, it plays an important role during the weight domain ontology construct [14] through validation of classification terms in their assigned taxonomy context.

Solution Agent: Allows us to filter and select from Universal Description Discovery and Integration (UDDI) registry the published solution or solved problems, which are available in the knowledge bases. This latter is important for our problem by applying a corresponding specific aggregation function. This functionality is not presented in the present paper.

Analyzer agent: This agent allows us to filter and select similar candidate problems [15] from deferent information source to construct our corpus and to assign it to the issue problem in order to generate our performance matrix via applying a set of algorithms and functions [16].

Meeting agent: Is responsible for preparing a collaborative session when no solution has been found trough the UDDI and in case where the problem is considered as incorrectly expressed or newly encountered as shown in the figure 4.
Other important components are presented in our architecture such as: the UDDI registry, Tree Tagger Dictionary, and knowledge base which are described as below:

Tree Tagger Dictionary: Is a tool for annotating text with part-of-speech and lemma information. In our approach, it is used during data pre-processing task as lemmatization tool, the latter one i.e. collects together the different forms of the same word.
Global Knowledge Bases: It includes the various interventions and scenarios that are used as similarity for the posed problems. It is the global data base whose information is accessible to all agents.
UDDI: It is an independent framework platform for describing, discovering, and integrating services. It provides a single location where services agents can dynamically find the stored solutions in UDDI which are published by different organizations.
Domain Ontology: Is developed to calculate the assigned weights to each criterion in our associated decisional performance table, and it used in measure similarity task.
The figure1 is an overview of a described architecture with some numbered flows to illustrate actions sequencing and relationships between all components that’s contain.

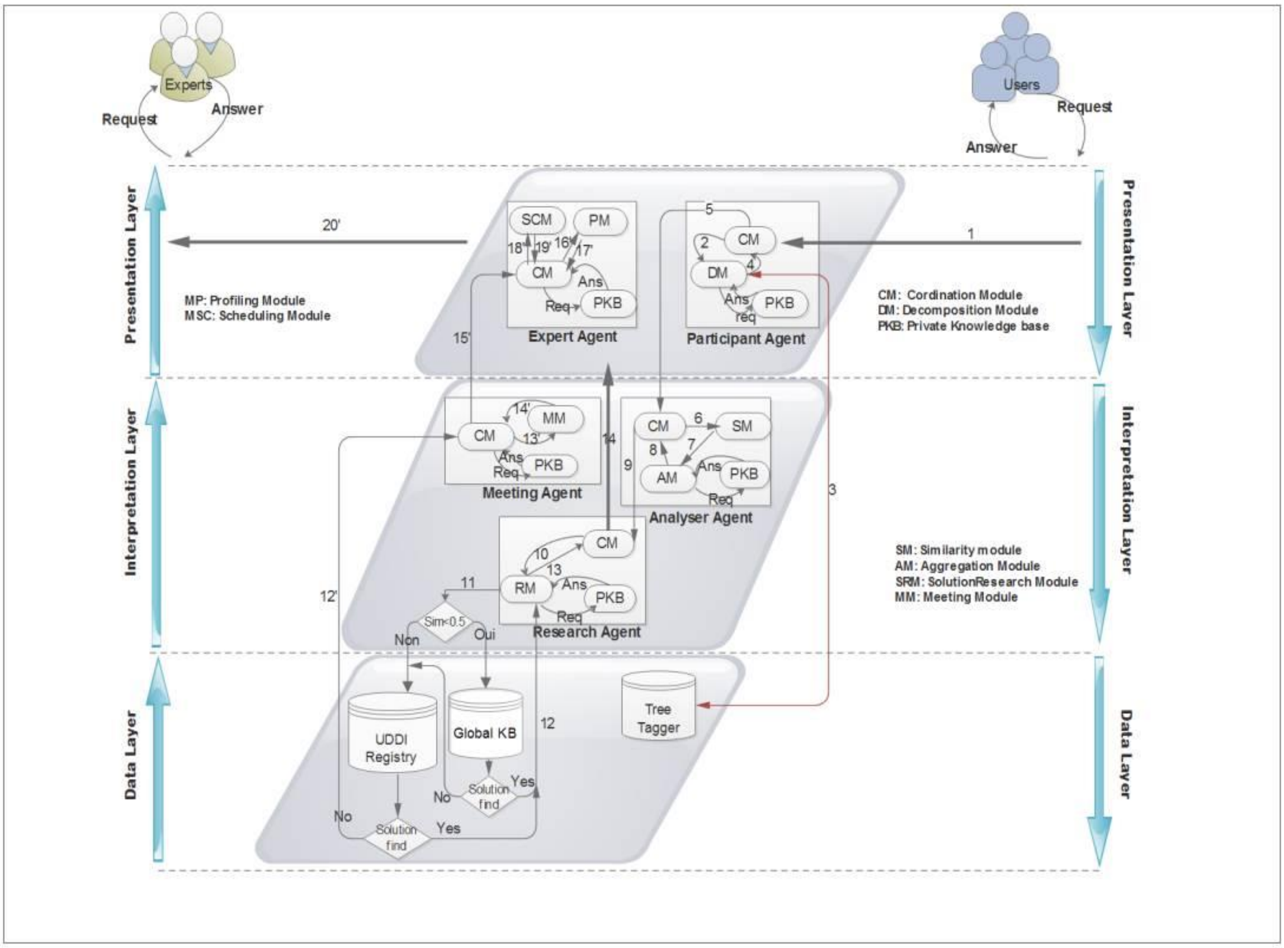


Figure 1: Multi Layers Agents based Architecture

This multi-layer architecture provides an ease of maintenance. Indeed, it allows monitoring of posed problem life cycle. However, a problem is considered solved if it contributes in all defined layers (layer1, layer2, and layer3). Moreover, it adds reliability and more independence to the involved agents and it gives us the ability to update the components of one layer without impacting the other layers. Indeed, to implement this we use Java NetBeans environment and Jade Platform for algorithms implementation and multi-agent deployment. In addition, Wamp server for data base deployment, Tree Tagger as dictionary and Protégé tools to develop domain ontology. In the following, we will describe the functionality of some modules that fulfill the structure of our agents:

Private knowledge Base (PKB): It contains knowledge that refers to tasks like processing function, aggregation methods parameters, and indexes on the shared domain ontology.
Similarity Module (SM): It resides with the expert agent in the fact to search and select all interested cases in the shared knowledge base using text similarity measure.
Aggregation Module (AM): is a part of analyzer agent task, is used in order to have an efficient solution rank with implement a specific function introduced in [16].
Coordinate Module (CM): This module plays an important role to control the whole mechanism of decision-making by coordinating between all agents that constitute the tree layers of our decentralized architecture in order to achieve the intended goal.

Decomposition Module (DM): Is responsible to generate all pertinent terms for each problem which are necessary to optimize and to accelerate the information retrieval in order to obtain

the best solutions.
Other modules are present in our approach such as profile module that resume overall expert profile and his preference. In addition the scheduling module that aims to organize a collaborative sessions with take account experts availabilities information's.

ILUSTRATION EXAMPLE

This section describes some interface of the **MAITD** defined in the latest section. Our corpus as a source information is constructed from real IT infrastructure data base problems.
Description Problems (Layer1): As first step, the participant describes a problem in his specific private space as shown in the figure2.

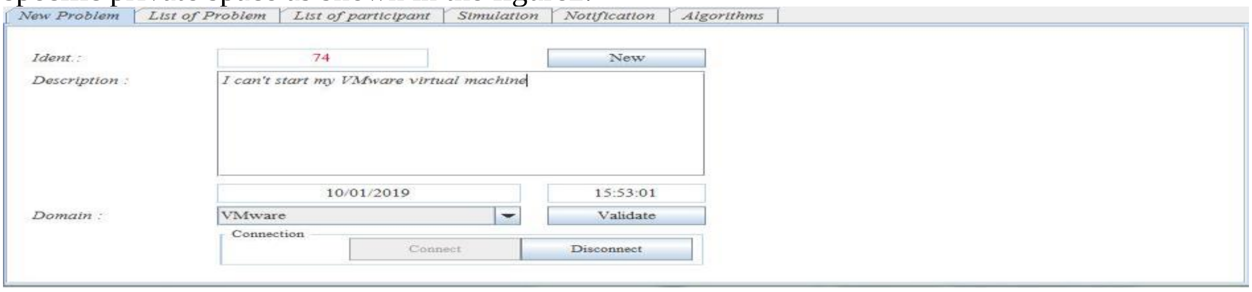
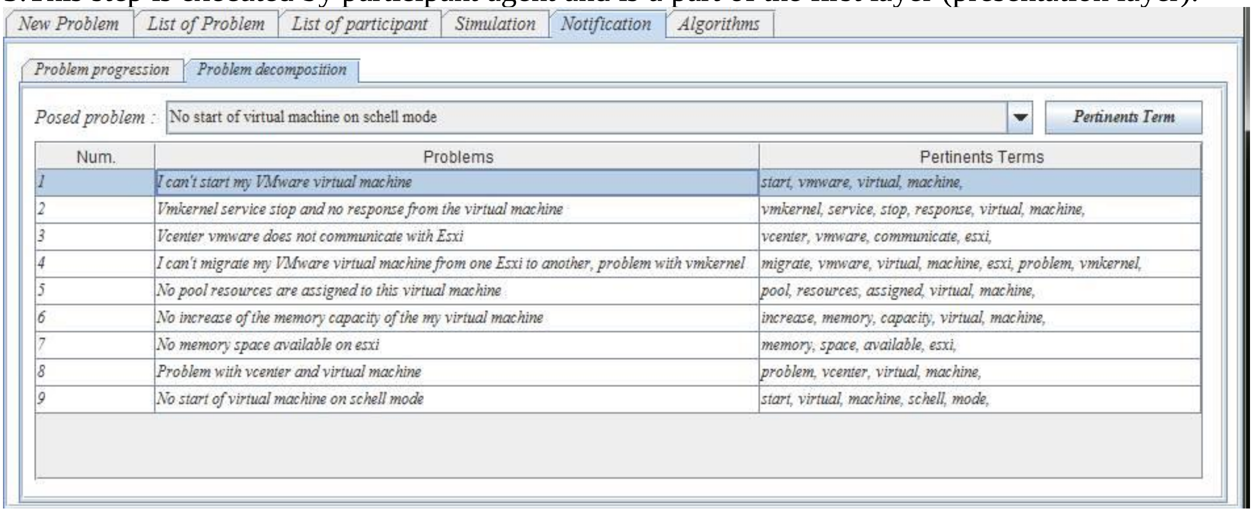


Figure 2: Problem Description Screenshot

Data pre-processing (Layer1):
Before implementing any method, it is initially necessary to transform the problems into an efficient manner, so that they can be analyzed. In this second step, we present the data after being formatted by using an adequate way to be dealt with by next phase as shown in the figure 3.This step is executed by participant agent and is a part of the first layer (presentation layer).



Num.	Problems	Pertinents Terms
1	I can't start my VMware virtual machine	start, vmware, virtual, machine,
2	Vmkernel service stop and no response from the virtual machine	vmkernel, service, stop, response, virtual, machine,
3	Vcenter vmware does not communicate with Esxi	vcenter, vmware, communicate, esxi,
4	I can't migrate my VMware virtual machine from one Esxi to another, problem with vmkernel	migrate, vmware, virtual, machine, esxi, problem, vmkernel,
5	No pool resources are assigned to this virtual machine	pool, resources, assigned, virtual, machine,
6	No increase of the memory capacity of the my virtual machine	increase, memory, capacity, virtual, machine,
7	No memory space available on esxi	memory, space, available, esxi,
8	Problem with vcenter and virtual machine	problem, vcenter, virtual, machine,
9	No start of virtual machine on schell mode	start, virtual, machine, schell, mode,

Figure 3: Data Pre-processing Screenshot

Preparing collaborative session (Layer2):
The meeting agents intervene to prepare collaboration sessions. Basically, there are two important points to note here. First, the meeting agent collaborates with all participant agents and expert agents to determinate this session. Second, this session will be scheduled if no solution is generated.

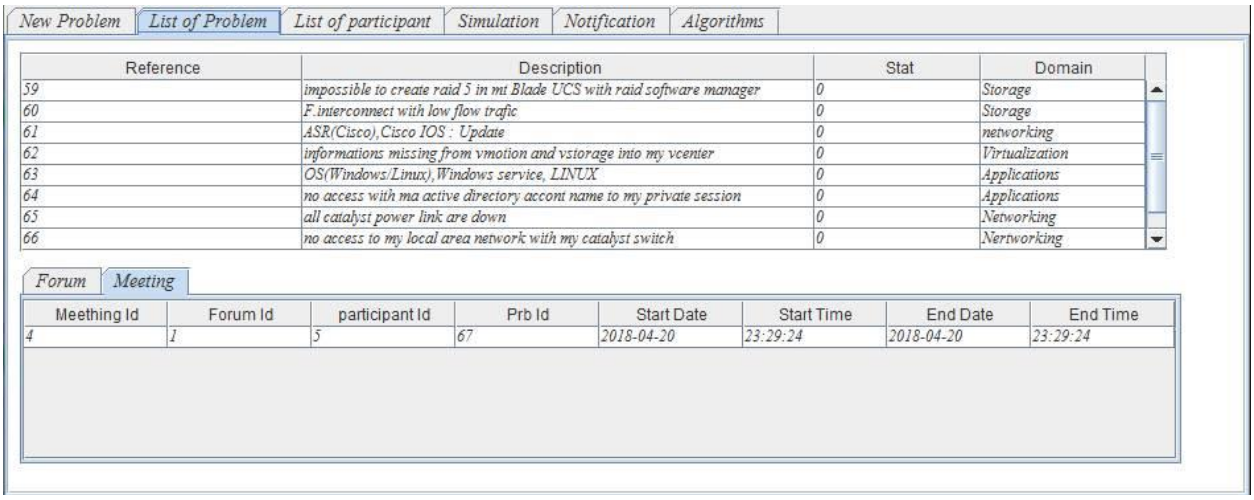


Figure 4: Preparing Collaborative Session Screenshot

Monitoring Problem Possessing (Multi Layers):

The participants can monitor all problems progress in the same interface and the different tasks performed by the agent's trough their private spaces as shown in the figure5.

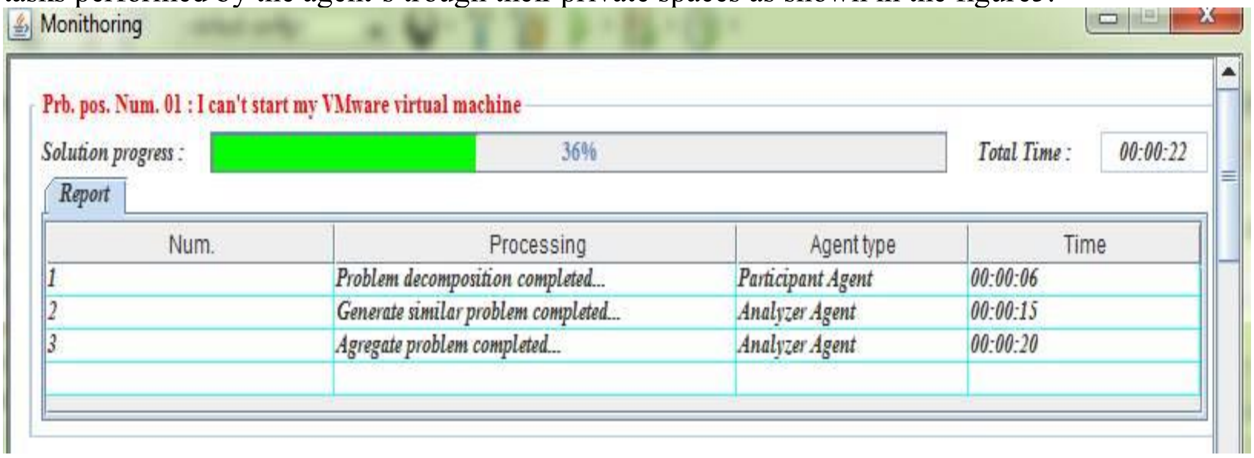


Figure 5: Participant Monitoring Screenshot

CONCLUSION

This paper presents multi-layer architecture for **MAITD** in which a decision support framework has been proposed to provide the expression and enforcement of corrective operational maintenance constraints among a set of agents that are allowed to communicate, coordinate, and cooperate to meet the goals of their particular users. In addition, a multi agent system aims to seek the best solution by taking into account the dynamic representation of the treated problem and their complexity. Moreover, it is able to analyze the traces of all agents in a continuous way (corrective monitoring) and it is also capable to react independently to the problem owner. For future work, we are going to improve our system, so that it moves from corrective maintenance to a preventive one. On the other hand, we will also show how the domain ontology is used to calculate the criteria weight and how the solutions are generated from the UDDI. Also we will discuss how similarity methods and solution aggregation functions are involved and implemented in our approach.

REFERENCES

[1] X. Mehmeti, B. Mehmeti, R. Sejdiu, "The equipment maintenance management in

manufacturing enterprises", PapersOnLine, 51-30 at www.sciencedirect.com ,2018.
[2] M.D. Kent, O. Costello, S. Phelan, K. Petrov, "Cost Oriented Maintenance management Systems for Manufacturing Processes", Waterford Institute of Technology, Ireland, 2017.
[3] P. Callewaerta , W.J.C. Verhagena , R. Curran, "Integrating maintenance work progress monitoring into aircraft maintenance planning decision support", 6th CEAS Air and Space Conference, Bucharest, 2017.
[4] Solar winds IT management and monitoring product, URL. https://www.solarwinds.com, Mars 2010.
[5] Manage Engine IT management and monitoring product, URL. https:// www.manage engine.com/ ,November 2007.
[6] Zenoss IT Infrastructure management and monitoring tool for hybrid IT environment, URL, https://www.zenoss.com/ , June 2018.
[7] H.J Liao , C.H.R Lin,Y.C Lin, K.Y Tung , "Intrusion detection system: A comprehensive review", Journal of Network and Computer Applications 36 , pages 16–24, 2013.
[8] Jaime Campos Om Prakash " information and communication technologies in condition monitoring and maintenance", 12th IFAC symposium on information control problems in manufacturing, Volume 39, Pages 3-8, 2006.
[9] J.N. Haack, G.A. Fink, W.M Maiden, D. McKinnon, S.J. Templeton, "Ant-based cyber security" , Proceedings - 2011 8th International Conference on Information Technology , ITNG, pp. 918–926, 2011.
[10] P. Kendrickm, N. Criado, A. Hussain, M. Randles, " A self-organising multi-agent system for decentralized forensic investigations", Expert Systems with Applications Volume 102, Pages 12-26, 2018.
[11] A. Jahanbin, A. Ghafarian, S.A.H. Seno, S. Nikookar , "A computer forensics approach based on autonomous intelligent multi-agent system" , International Journal of Database Theory and Application, pp. 1-126 , 2013.
[12] D. Rudrapal, S. Das, N. Debbarm, S. Ebbarma, "Internal Attacker Detection by Analyzing User Keystroke Credential", Lecture Notes on Software Engineering vol. 1, no. 1, pp. 49-52, 2013.
[13] S. Elandaloussi , N. Taghezout, P. Zaraté, "A Collaborative Solution for IT Infrastructure Maintenance Based on Web Services and Mobiles Agents", Namur, Belgium, 2017.
[14] E. R. Bendaoud, A.M.R Hacene, Y. Toussaint, B. Delecroix, A. Napoli, "Construction d'une ontologie à partir d'un corpus de textes avec l'ACF", 18eme journées francophones d'ingénieries de connaissances, Grenoble, 2007.
[15] A. islam, D. Inkpen, "Semantic text similarity using corpus based word similarity and string similarity", Transaction on knowledge discovery from data, New York, USA , 2008.
[16] S. Fomba, "Un système de recommandation pour le choix de l'opérateur d'agrégation", PhD, Toulouse University, Institut de recherche en informatique de Toulouse, July 2018.

Social Web Mining as a Tool to Support Public Security Sentiment Analysis

Victor Diogho Heuer de Carvalho[1,2], Ana Paula Cabral Seixas Costa[1]

[1]Universidade Federal de Pernambuco
Av. da Arquitetura, s/n, Cidade Universitária,
Recife - PE, 50740-550
victor.hcarvalho@ufpe.br, apcabral@cdsid.org.br
web-page: http://www.ufpe.br/dep

[2]Universidade Federal de Alagoas
Rodovia AL 145, Km 3, nº 3849, Cidade Universitária,
Delmiro Gouveia – AL, 57480-000
victor.carvalho@delmiro.ufal.br
web-page: http://www.ufal.edu.br/sertao

ABSTRACT

This short paper provides a preliminary overview of ongoing research to develop, test, and validate an instrument for decision support, using social web mining and sentiment analysis, with its primary application in the area of Public Security management. The paper presents fundamental concepts collected from an initial survey of the literature, to assist in justifying the proposed framework. The Design Science in Information Systems guides the methodology. It helps to identify objectives and describe elements necessary to guarantee to this specialized community the qualification of the research and the reliability of the final product. The proposal means to determine the general elements that make up the framework, highlighting some techniques for: (i) extraction and storage of social-web user records; (ii) preprocessing and storage of unstructured data; (iii) data processing with noteworthy use of machine-learning algorithms; (iv) sentiment analysis using natural language processing, to eliminate biases of vagueness or irony; and (v) visual presentation of results. As the main intended result, the framework should at least support public managers in defining strategies aligned with public opinion on local security.

Keywords: Public security management, Framework for decision support, Unstructured data, Social web mining, Sentiment analysis.

INTRODUCTION

The search for appropriate ways to retrieve information from a wide variety of sources and in increasing amounts has led organizations to use concepts related to unstructured data mining. In this context, this paper proposes text mining and, more specifically, social-web mining as means of finding patterns or trends with the power to explain a certain phenomenon [1].

Above all, both text mining and social-web mining are processes of knowledge discovery applied respectively to textual databases [2] and registers (mainly textual) of users of social networks [3]. Inserted in this set is sentiment mining or analysis that aims to study patterns of feelings, opinions, emotions, and behaviors expressed through texts, making use of computational tools [4] such as Natural Language Processing [5].

Among these concepts, Big Data is another emerging concept in the context of organizations

seeking to supply themselves with data from a variety of sources for use in their decision-making process [6]. Modern models of decision-support systems consider analytical tools that go beyond traditional data mining (of structured data), incorporating data from a variety of sources and following the Big Data trend [7].
Several organizational and social applications utilize these technologies, from the identification of behavioral tendencies of groups of individuals related to products and brands, through the understanding of economic behavior, to trends in politics and opinions on the most diverse types of phenomena affecting social welfare [3, 5, 8].
From this conceptual set emerges the idea that this short paper presents: an initial report on an ongoing study of the application of social-web mining for the purpose of supplying data that supports the sentiment analysis of public safety in a large Brazilian city. Following is an initial plan for methodological procedures, as well as the initial general idea of the framework necessary for the realizing the mining and analysis processes, as well as some concluding remarks.

RESEARCH METHODOLOGY

The research follows the Design Science methodology applied to Information Systems [9], since its goal is to formulate a framework for decision-support systems for public security policies, with social-network mining tools for public managers to use. For this purpose, with the methodology so defined, alignment of organizational needs in the application area in question requires identifying available technologies and how they can act strategically to support the necessary decisions.

Strictly following this methodology enables determination of a basis with the necessary guidelines, so that the main artifact developed will align with the expectations of its end users (in terms of requirements), and be tested, validated, and communicated to the technical-scientific community, guaranteeing its reliability. Table 1 summarizes the guidelines according to the Design Science methodology applied to Information Systems.

Table 1: Synthesis of the methodological guidelines for the research.

Guideline	Description / Purpose
1st Design as an artifact	To propose a framework for decision support systems in public security policies based on social-web mining and sentiment analysis.
2nd Problem relevance	Managers (decision makers) in the general area of public management, need objective tools that present them with structured and intelligible data to support their decision-making processes.
3rd Design evaluation	In order to corroborate the relevance of the problem, formal research methods will be sought for specific elements to carry out the related evaluations.
4th Research contributions	The framework, considered the final product of the research, should be effectively replicable not only in the area of management and research target, but in any environment with similar needs.
5th Research rigor	The scientific method should be rigorously applied to ensure that the results are consistent and the product is reliable, satisfying the explicit needs of the users for whom it is intended.
6th Design as a search process	Search in the scientific literature of the areas involved, cases already studied, existing models, consistent design/engineering of artifacts, with foundation in the concepts involved in order to support the final idea.
7th Communication of research	Disseminate to the area scientific and technical community the progress of the research, to obtain feedback and opinions about the framework that will aid in its improvement.

Figure 1 shows a general flowchart of the search to be performed, based on the previously determined guidelines.

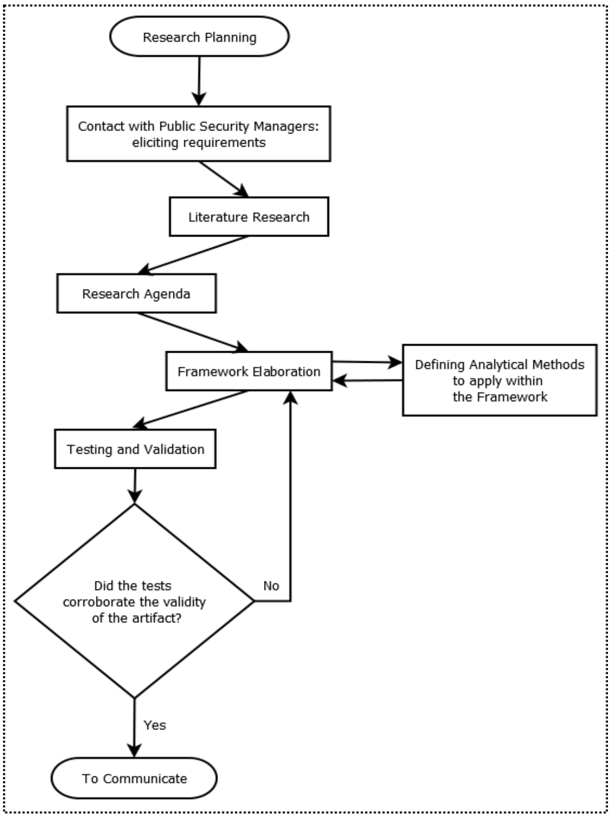


Figure 1: General research flow.

Notably, the flowchart shows that contact with the managers of the target area, together with the literature review, will provide requirements and a set of formal elements for the process of elaborating the framework, which should follow a research agenda. Testing and validation will be fundamental to corroborating the validity of the product, and once again participation of the area managers is indispensable. In this step, as to whether the decision on the artifact is valid, cycles will guarantee the improvement of the idea. At the end of a cycle with proper validation, the results will be communicated to the community (with publication of articles and participation in conferences).

Given this presentation of the applicable methodological foundations, following are some preliminary results related to the main product being sought.

PRELIMINARY DEFINITIONS OF A FRAMEWORK FOR SOCIAL-WEB AND SENTIMENT MINING

The essential data sources in the framework that ongoing research seeks are large existing social networks, focusing primarily on Twitter, and then on Facebook, through their application programming interfaces (APIs), for the application of scripts containing data-extraction routines according to queries structured with key terms. A similar process is established [10] through a knowledge-management framework integrating the concepts and tools of Big Data and data mining in social media. Another work [11] that proposes frameworks for analysis of social-web content has the objective of performing the classification of tweets through a set of specific rules with distinctly different approaches.

Having retrieved the necessary data from their sources, special treatment techniques must be applied, creating a preprocessing phase where a kind of data cleaning occurs, for further processing and sentiment analysis. The preprocessing phase contains methods such as normalization, tokenization, and filtering and extraction of features [12], which will treat the data collected through the text records, leaving them a standard fit for the final processing that will come later, opening the way for the sentiment analysis.

In this phase of analysis of feelings, several algorithms can be applied, with the initial option the set referring to machine learning, where there are supervised and unsupervised algorithms. Among the most famous algorithms are the decision trees, the support vectoring machines, the Bayes naïve classifier, Bayes networks, neural networks, and the maximum entropy method [13].

Figure 2 presents a preliminary schema of what the framework will involve. Because it is based on an initial idea, its structure must undergo changes that will reflect what the research detects, going from a generalist standard to something much more specialized.

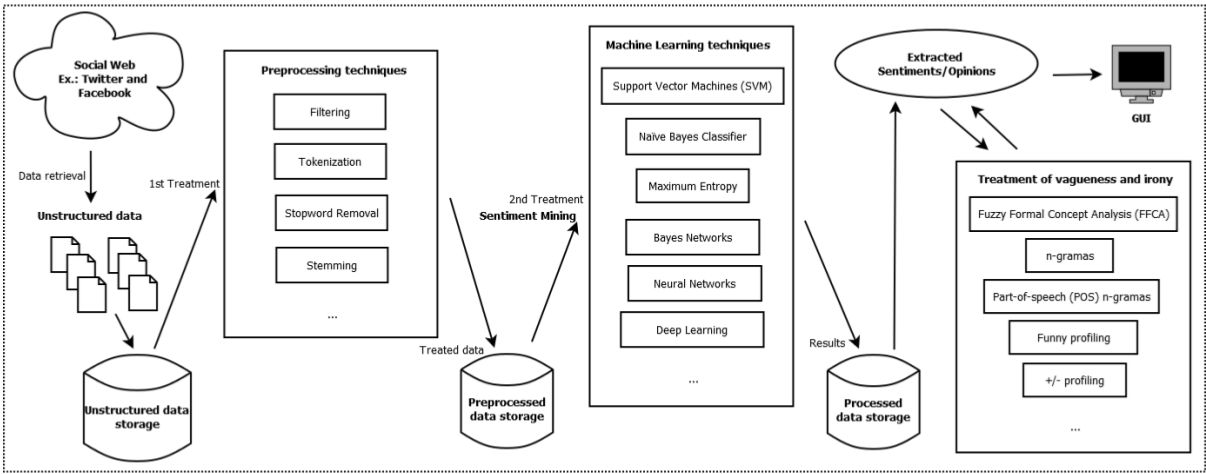


Figure 2: Preliminary framework conception.

As the schematic shows, a final stage of applying vagueness and irony treatments should eliminate biases in data interpretation and provide a clearance for the processing done previously. Development of a graphical user interface (GUI) will ensure that end users (public managers) can select various ways of presenting results.

The basis for the implementation of the analyses contained within the framework will be the general-purpose programming language Python. The choice is justified first because it is a free language, without costs involved; then by virtue of the power it possesses, given its extensive set of libraries dedicated to the extraction and analysis of data (whether structured or not). Some examples are Tweepy and Python-Twitter, to work with the Twitter API; Facebook, for the case of the API of the social network with the same name; Natural Language Toolkit (NLTK) to handle text processing and mining and sentiment/opinion analysis; Scikit-Learn, for machine learning, among countless others. The GUI can also be developed using a framework for object-oriented web systems in Python language: Django.

APPLICATIONS AND CONCLUDING REMARKS

The application to which the research is initially oriented is the Public Security area, focusing on the collection of data (initially on Twitter) about the topic from among users delimited to the metropolitan region of Recife, capital of the state of Pernambuco (Brazil). As a fundamental analytical result, it is intended to analyze the opinion or feeling of the users regarding public security in their respective places of residence, based on the retrieved tweets, for certain time bands and within the geographical limits of the region. The tweets will be extracted in the format "json" (JavaScript Object Notation), specifically for data and metadata transport, from which analysis will produce some of the most relevant tags for the case (referents to the dates, georeferencing data, and the text of the user's post).

This essential data will enable a spatial and temporal analysis of the public user's sentiments/opinions about public security, providing through a specialized GUI information on (for example) regions with a higher incidence of complaints (i.e., negative sentiments) on issues of safety, allowing in a second step the execution of studies correlating these sentiments/opinions with such quantitative data as police units, vehicles, and registered occurrences of crimes.

Finally, the possible applications of the framework will not be restricted to only the domain of public security. The author-researchers' intention is to reach other key areas of public infrastructure, such as education, health, and transportation, after validation in the initial area.

FOUNDING

This work was financed in part by the Coordination for the Improvement of Higher Education Personnel (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, Brazil) – Finance Code 001, and by the National Council for Scientific and Technological Development (Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, Brazil).

REFERENCES

[1] Z. Zainol, M. T. H. Jaymes, and P. N. E. Nohuddin, "VisualUrText: A Text Analytics Tool for Unstructured Textual Data," *Journal of Physics: Conference Series*, vol. 1018, p. 012011, May 2018.

[2] A. Usai, M. Pironti, M. Mital, and C. A. Mejri, "Knowledge discovery out of text data: a systematic review via text mining," *Journal of Knowledge Management*, vol. 22, no. 7, pp. 1471–1488, 2018.

[3] A. Sapountzi and K. E. Psannis, "Social networking data analysis tools & challenges," *Future Generation Computer Systems*, vol. 86, pp. 893–913, 2018.

[4] K. Ravi and V. Ravi, "A survey on opinion mining and sentiment analysis: Tasks, approaches and applications," *Knowledge-Based Systems*, vol. 89, pp. 14–46, 2015.

[5] S. Bjurstrom, "Sentiment Analysis Methodology for Social Web Intelligence," in *Proceedings of the Twenty-first Americas Conference on Information Systems*, pp. 1–12. AIS, 2015.

[6] M. Kowalczyk and P. Buxmann, "Big Data and Information Processing in Organizational Decision Processes," *Business & Information Systems Engineering*, vol. 6, no. 5, pp. 267–278, Aug. 2014.

[7] T. Poleto, V. D. H. Carvalho, and A. P. C. S. Costa, "The Full Knowledge of Big Data in the Integration of Inter-Organizational Information," *International Journal of Decision Support System Technology*, vol. 9, no. 1, pp. 16–31, Jan. 2017.

[8] D. S. Sisodia and N. R. Reddy, "Sentiment analysis of prospective buyers of mega

online sale using tweets," in *Proceedings of the IEEE International Conference on Power, Control, Signals and Instrumentation Engineering of 2017*, pp. 2734–2739. IEEE, 2018.

[9] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design Science in Information Systems Research," *MIS Quarterly*, vol. 28, no. 1, p. 75, Dec. 2004.

[10] W. He, F. K. Wang, and V. Akula, "Managing extracted knowledge from big social media data for business decision making," *Journal of Knowledge Management*, vol. 21, no. 2, pp. 275–294, 2017.

[11] M. Z. Asghar, F. M. Kundi, S. Ahmad, A. Khan, and F. Khan, "T-SAF: Twitter sentiment analysis framework using a hybrid classification scheme," *Expert Systems*, vol. 35, no. 1, pp. 1–19, 2018.

[12] H. Al-Saif and H. Al-Dossari, "Detecting and Classifying Crimes from Arabic Twitter Posts using Text Mining Techniques," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 10, pp. 377–387, 2018.

[13] Z. Song and J. Xia, "Spatial and Temporal Sentiment Analysis of Twitter data," in C. Capineri, M. Haklay, H. Huang, V. Antoniou, J. Kettunen, F. Ostermann and R. Purves, (eds.), *European Handbook of Crowdsourced Geographic Information*, London: Ubiquity Press, pp. 205–221, 2016.

A Conceptual Model based on Normative Multiagent System for Supply Chain Integration of Brazilian Gypsum LPA

Thárcylla Rebecca Negreiros Clemente, Aldênia Karla Barrêto Candido
Management Engineering CAA / Universidade Federal de Pernambuco
Av. Campina Grande, s/n - Km 59 - Nova Caruaru, Caruaru - PE, 55014-900, Brazil
thnegreiros@ymail.com, aldeniak@gmail.com

ABSTRACT

The Brazilian gypsum has one of the highest levels of recognized mineral quality in the world. Companies that operate in this sector support major economy-driving industries such as civil construction, architecture, medicine and agriculture and are an input for the manufacture of paints, phosphors and insecticides, among others. Most of the companies that constitute this sector are small and medium-sized enterprises (SMEs), which require supply chain solutions to maintain competition in market share. Supply chain integration as a collaborative solution has become a necessary strategy for enhancing SMEs decision-making capability. This study proposes a conceptual model based on a Normative Multiagent System Approach that offers a dynamic view about supply chain integration for managers and others stakeholders. From this contribution, it is possible to obtain constant information flow among stakeholders and build an effective strategic context for SMEs in the Brazilian Gypsum Local Productive Arrangement (LPA).

Keywords: Normative Multiagent System, Supply Chain Integration, Local Productive Arrangement, SMEs, Brazilian Gypsum Industry.

INTRODUCTION

Nowadays, even with advances and technological innovations, gypsum remains a mineral widely used in construction, architecture, orthopedic and dental medicine and agriculture and as an input for the manufacture of paints, phosphors and insecticides, among others [1]. In Brazil, the exploration and commercialization of gypsum, a raw material are capable of economically boosting the region of Araripe, whose productive and economic activities involving gypsum are governed by Gypsum Local Productive Arrangement (LPA) of Pernambuco [2, 3]. This LPA, which has the best quality ore in the world, has an abundant reserve of gypsum, besides presenting excellent conditions for exploration of the material and geomorphology of the deposit.

SMEs that compose the Gypsum LPA are encouraged to effectively manage all their processes, this being possible through the development of strategies and management on the flows of information and products, so that they can efficiently traverse the entire supply chain. Supply Chain Management (SCM) is a management model that requires organizational changes in order to develop cooperation, communication and trust among all chain stakeholders [4]. These changes are related to the integration of cultural, operational and technological aspects considered by the stakeholders as main obstacles inserted into structural characteristics, conduct and market performance.

The stakeholder's autonomy on decision making requires that the norms, roles and responsibilities of each institution are defined. The rules define the regulated contributions of the behavior of each institution as an agent of the supply chain. In this context, the public policies of the Brazilian government play an essential role in defining the guidelines for the mineral extraction market, and its inclusion as regulator of the supply chain integration process is fundamental. From this perspective, the present work presents a conceptual model for the integration of the Gypsum LPA using the normative multiagent systems approach, considering the perspectives of communication, interaction and information flow to support the decision-making process of the main stakeholders.

Brazilian Gypsum LPA

LPA can be understood as a territorial agglomeration with a presence of economic, political and social agents that are involved in a constant transformation and a specific set of activities with common interests that communicate with each other. In general, LPA are formed by SMEs that exert a major influence on the competitive market.
Gypsum LPA of Pernambuco is made up of SMEs in industries such as mineral extraction, calciner and/or artifacts. This LPA offers great economic potential for a region since this is the largest set of deposits exploited in the country, representing 95% of the plaster demanded nationally [2]. The industry supply chain comprises several entities that take responsibility for the flow of products and information for market integration. Table 1 presents the main components of the supply chain integration (SCI) process of plaster LPA.

Table 3: Supply Chain Integration Context

Supply Chain Requirements	Description
Main actors/stakeholders	Customers, retailers, wholesalers/distributors, manufacturers, component/raw material suppliers, society, trade union, financial agencies, educational institutions, business consulting agencies, and government.
General objective	Actors are responsible for the increase of the companies' market share and for improving customer expectations of the supply chain.
Transformation process	Transformation of gypsum is the main aim of the companies. The outputs are finished products for several areas: health, agriculture, construction.
Enviromental constraints	There are barriers set that the system can face such as economic situations, scarce raw materials and substitute products, among others.

Actors are stakeholders or agents, individuals and target groups that cooperate, voluntarily or not, to create wealth within organizations, thus being the beneficiaries and/or potential risk people, and they exert an important role in ensuring that cooperation is efficiently achieved in the SCI process. In this sense, it indicates organizations that manage to treat their stakeholders generously achieve higher levels of competitiveness by increasing company values. From this, they tend to be reciprocal in their attitudes; for example, the stakeholder-client can increase requests, the stakeholder-society can increase tax incentives and the stakeholder-financial institutions can facilitate fundraising, among other examples. The influence of each agent in this process encourages the investment of the flow of information that helps in the development of the business, due to the cooperation and integration of shared knowledge, and aids in the decision-making of the SCI process.

DECISION-MAKING IN SUPPLY CHAIN INTEGRATION PROCESS

The SCI process enables improving the coordination of activities and information flow to support the decision-making process among organizations, providing greater competitive potential, greater knowledge related to the power of communication and the transfer of information among stakeholders [5, 6], especially for SMEs which must face cultural, operational and technological obstacles [6]. These obstacles encourage studies on management alternatives that minimize impacts on the decision-making process.
In the supply chain, companies have several control points distributed in a dynamic environment, and each one has a specific objective, but they interrelate in order to solve tasks to cooperate in the market. This means that the collaborative relationships are related to the profits gained by each agent as a member of the chain and by the whole chain [7]. Figure 1 shows a supply chain decision-making framework based on logistical and cross-functional drivers [8].

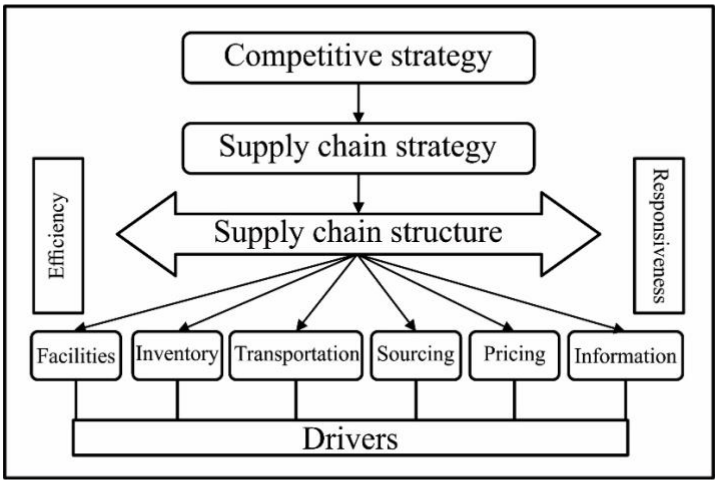


Figure 10: Supply Chain Decision Making Framework [8]

Figure 1 shows the main management drivers: (i) Facilities, Inventory and Transportation are logistical drivers that make decisions about the assets that a company owns or outsources; (ii) Sourcing, Pricing and Information are cross-functional drivers that make decisions about how each logistical driver will be obtained, used, delivered and so on. All actions should be provided in the supply chain strategy to answer to production capacity, demand forecasting and customer satisfaction, among others. These initiatives can be used to ensure market share, active communication, and compliance with rules and laws imposed by market regulatory institutions.
In this context, there are public policies that define laws and regulations that ensure the economic incentive and regularization of productive activities carried out by companies. The legislative aspects are fundamental to defining the behavior of the agents and should be incorporated into the conceptual model for the SCI process.

NORMATIVE MULTIAGENT SYSTEM

The normative multiagent system definition can be understood as combined concepts about normative systems and multiagent systems [9]. Norms are essential to conducting human social behavior and are fundamental elements to describe artificial agents' performance that simulate and collaborate with humans. Norms are implemented in multiagent systems to provide models for cooperation and co-ordination among artificial agents that represent the human intelligence rationale in social contexts.

By integrating norms and multiagent systems, it is possible to build and implement normative multiagent systems with specific instructions to solve group decision problems, regulate societies and structure multiagent organizations, among other cases [9]. A normative system is identified if there is at least one normative correlation among a set of sentences that result in normative consequences in specific cases [9]. The agents interact with the normative system and are able to create new norms, update or maintain norms and enforce norms [10] using regulation rules, legislators or other government forces, and they can use those rules to determine the constraints and conduct equilibrium of social behavior. In normative multiagent systems, the agents use the norms to define their autonomy as decision makers in a dynamic system.

NORMATIVE ASPECT FOR BRAZILIAN GYPSUM LPA

For SMEs in the Brazilian Gypsum LPA context, there is strong influence of normative instructions such as laws and regulations, which are guidelines imposed by the public power in order to guide the actions, rules and procedures for the relationship and behavior of the main actors in this market. Among the guidelines, the competitive strategies are related to the economic model, technological innovations, productive restructuring and the effects of job creation and reduction of social inequality in the local region.

Decision-making processes require an active flow of information to ensure control and interaction among different agents in the market. Figure 2 represents the role of normative agents highlighting the flow of information developed for the Brazilian Gypsum LPA context.

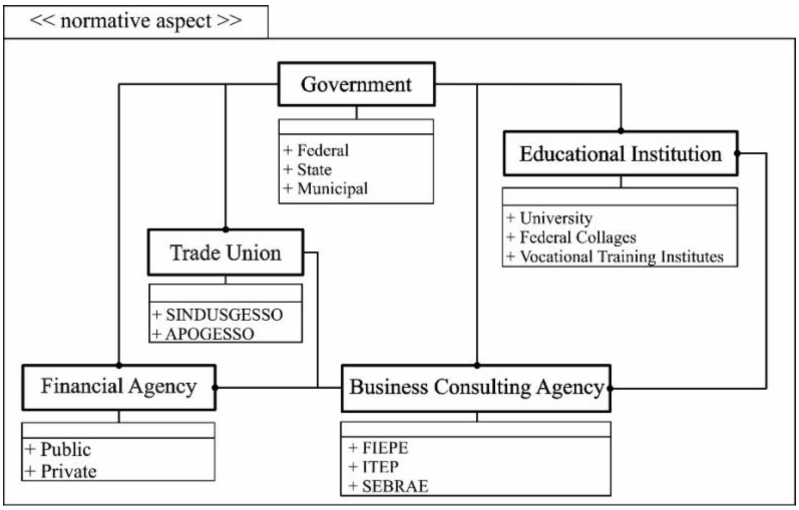


Figure 2: Normative Agents in Brazilian Gypsum LPA

In general, the actors who develop the guidelines are represented by educational institutions, business consulting agencies, trade unions, financial agencies and government. As shown in Figure 2, the government agent provides basic infrastructures for educational institutions and exerts influence on all the other agents. Each of the agents assumes their autonomy, but they act in the same dynamic environment. This relation is associated with public power, which define the roles of the financial agencies, trade unions and business consulting agencies. These agents influence the companies in the sector, especially when defining tax incentives, the legalization of informal enterprises, the provision of managerial and technological knowledge and financial support from LPA to continue generating revenue and manpower absorption. From this, the normative aspect is essential to describing the SCI process.

Conceptual Model for Supply Chain Integration

Based on Figure 2, the multiagent systems approach is able to represent cognitively active agents in the same cooperation environment. In this environment, the roles and responsibilities of each agent are defined and maintain the constant flow of information to ensure the best performance of agents. Figure 3 shows a diagram of the SCI process considering SMEs in the context of Brazilian Gypsum LPA.

1.

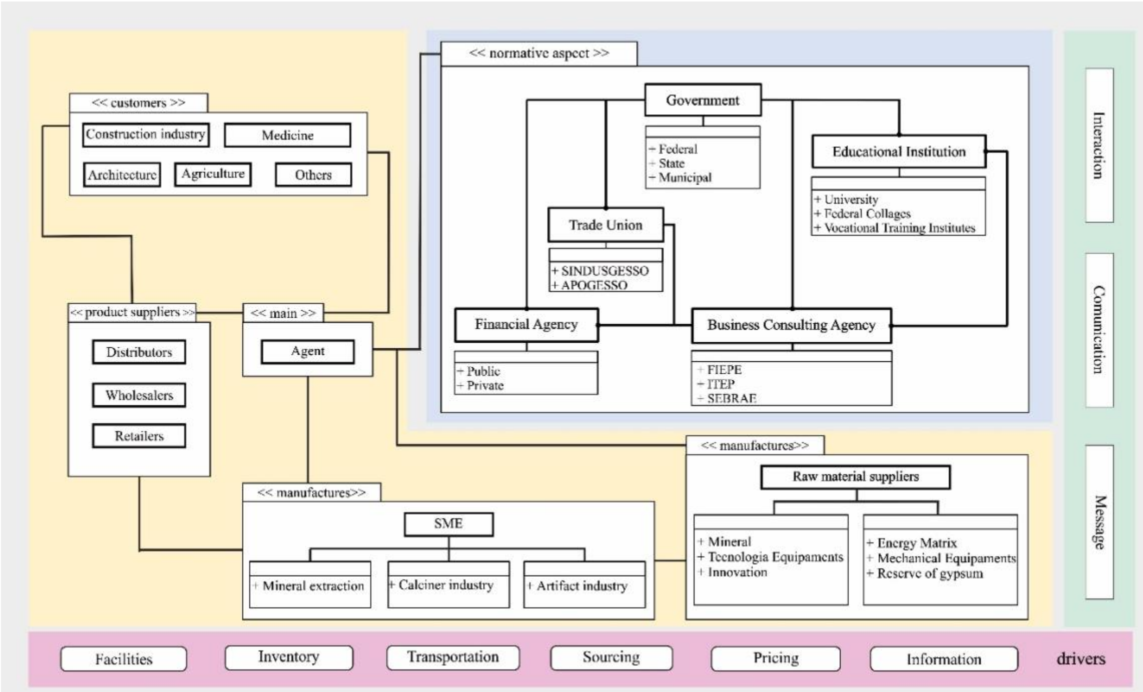


Figure 3: UML Diagram for SME Supply Chain Integration

As illustrated in Figure 3, the provision of models, tools and collaborative technologies will enable a more effective business strategy for handling the products and agents service perspective of key SMEs in the SCI process. In this case, the supply chain structure is represented by customers, product suppliers, manufacturers, raw material suppliers and normative agents. Raw material suppliers provide active and passive resources to the manufacturers, which provide finished or semi-finished products to retailers, wholesalers and/or distributors. The product suppliers deliver the output to the different customers. All agents assume a specific role determined by normative aspects.

The functional requirements in the conceptual model correspond to the interaction, communication and sending of messages to support the decision-making process based on logistical and cross-functional drivers. Thus, this conceptual model presents an adequate contribution for the SCI process in the Gypsum LPA context.

CONCLUSIONS

The Brazilian Araripe region is responsible for an expressive volume of production and commercialization of gypsum and its derivatives throughout the nation' territory. Therefore, studies on the SCI process are significant in providing opportunities for internalization of regional public policies. From this work, a conceptual model based on the normative multiagent system approach represents the SCI process of the Brazilian Gypsum LPA, which includes normative instructions, such as norms, laws and regulations, about stakeholders' market share to define their behavior in this context.

This model for the SCI process is at the development stage. For this work, the initial results are satisfactory and offer an improvement regarding Brazilian Gypsum supply chain management that can be considered a collaborative solution for competitive strategy and for enhancing the decision-making process for SMEs in the LPA. Thus, the model will be used to support different decisions and to indicate the best recommendations for maintenance of interaction, communication and flow of information among stakeholders.

ACKNOWLEDGMENT

This work has been partially supported by FACEPE and CAPES.

REFERENCES

[1] DNPM. Brazilian Mineral Balance. Available on www.dnpm.gov.br, 2019.
[2] ITEP, Technology Institute of Pernambuco. Available on www.itep.br/arranjos-produtivos-locais, 2019.
[3] H. Lastres, J.E. Cassiolato, "Innovation systems and local productive arrangements: new strategies to promote the generation, acquisition and diffusion of knowledge". Innovation: Management Policy and Practice, 7(2/3), 172, 2005.
[4] M. Braunscheidel, N. Suresh, A. Boisnier, "Investigating the impact of organizational culture on supply chain integration", Human Resource Manag., 49(5), 883-911, 2010.
[5] C. Ataseven, A. Nair, "Assessment of supply chain integration and performance relationships: A meta-analytic investigation of the literature", International Journal of Production Economics, 185, 252-265, 2017.
[6] S. Palomero, R. Chalmeta, "A guide for supply chain integration in SMEs", Production Planning & Control, 25(5), 372-400, 2014.
[7] D. D. P. Perales, F.C. L. Esteban, M. M. E. A. Díaz, J. E. Hernández, "Framework for Modelling the Decision: View of the Supply Chains Collaborative Planning Process", International Journal of Decision Support System Technology, 4(2), 59-77, 2012.
[8] S. Chopra, P. Meindl, "Supply Chain Management: Strategy, Planning and Operation, 5th Edition, Pearson, 2013.
[9] G. Boella, L. van der Torre, H. Verhagen, "Introduction to normative multiagent systems", Comput Math Organiz Theory, 12, 71–79, 2006.
[10] Y. Shoham, M. Tennenholtz, "On the synthesis of useful social laws for artificial agent societies" (Preliminary report), In: Proceedings of the National Conference on Artificial Intelligence. San Jose, 276–281, 1992.

Special Talks

Aiding to Decide and Aiding to Design

Alexis Tsoukiàs
CNRS Research Director at LAMSADE, PSL University, Université Paris Dauphine
National Coordinator of the Policy Analytics Research Network

The talk explores the two topics (deciding and designing) and shows that many activities (and artifacts) developed in order to support the process of deciding and/or designing are similar if not the same. It also shows that these two processes are intimately connected since in order to improve decisions we need to improve designs and in order to improve designs we need to improve decisions. For this purpose, we introduce a new framework. The talk uses both simple examples from the OR and Decision Analysis literature as well as complex problem situations in order to support the claims.

On Multi-Criteria DSS dedicated to Environmental Evaluation and Sustainability Assessment

João Carlos Namorado Clímaco
Emeritus Professor at the Faculty of Economics of the University of Coimbra
Researcher & Advisory Board Member of the INESC-Coimbra Research Center
Editorial Board Member of the University of Coimbra Press.

The concept of sustainability development is still arguable. However, the major idea was established by the famous Brundtland Report characterizing it as “meeting the needs of the present without compromising the ability of future generations to meet their own needs”. Nowadays, sustainability is a multidimensional subject, involving social, economic and environmental conflicting issues and, in general, several actors. Furthermore, environmental evaluation is per se an intricate multicriteria-problem. So, the use of Multi-Criteria Decision Support Systems seem to be adequate. In the first part of the talk we make a critical state of the art review of this area. In the second part of the talk we discuss the characteristics of the most appropriate multi-criteria approaches to deal with these complex issues, involving deep uncertainty. The adequacy of DSS architectures is also discussed. It will be argued that learning oriented tools enabling to deal with incomplete information are the most appropriate. A methodological discussion on some specific decision support tools based on our research will be carried out and hints on further developments are presented. Some Brazilian case studies, carried out in collaboration with a team of UFRJ (federal University of Rio de Janeiro), will be used to illustrate our points of view.

Web-based Approaches for DSS

Logikós: Augmenting the web with multi-criteria decision support

Alejandro Fernandez, Gabriela Bosetti, Sergio Firmenich
LIFIA, CIC/Facultad de Informtica, UNLP. 50th St. and 120 St., 1900 La Plata, Argentina
{name.surname}@lifia.info.unlp.edu.ar

Pascale Zarate
IRIT, Toulouse Université, 2 rue du Doyen Gabriel Marty, 31042 Toulouse Cedex 9, France
{name.surname}@irit.fr

ABSTRACT

There are activities that on-line customers daily perform, which involve a multi-criteria decision challenge. Choosing a destination for traveling, buying a book to read, or buying a mobile phone are some examples. Customers analyze and compare alternatives considering a set of shared characteristics, and under certain criteria. E-commerce websites frequently present the information of products without special support to compare them by one or many properties. Moreover support for decision making is limited to sorting, filtering, and side-by-side comparison tables. Consequently, customers may have the feeling that the merchants interests influence their choices, which are no longer grounded on the rational arguments they would like to put in practice. Moreover, the alternatives of interest for the customer are frequently scattered across various shops, with no support to collect and compare them in a consistent and customized manner. In this article, we propose empowering users with multi-criteria decision making support on any website, and across different websites. We also present Logikós, a toolbox supporting multi-criteria decision making depending on the presentation layer of any Web page.

Keywords: E-commerce · Multi-criteria decision support · Web Augmentation

Find and cite the full work as:
Fernandez, A., Bosetti, G., Firmenich, S., Zarate, P., 2019. Logikós: Augmenting the web with multi-criteria decision support, in: Freitas, P.S.A., Dargam, F., Moreno, J.M. (Eds.), Decision Support Systems IX: DSS Developments & Future Trends, Lecture Notes in Business Information Processing. Springer International Publishing, Cham, pp. 116-128.

A multicriteria web-based Decision Support System to allocate resources in an electricity company using the c-optimal portfolio concept

Lucas Borges Leal da Silva[1], Adiel Teixeira de Almeida[2], Jonatas Araújo de Almeida[3]
CDSID – Center for Decision Systems and Information Development, Universidade Federal de Pernambuco
Avenida da Arquitetura, Recife 50740-550, Brasil
[1]borgesleal.lucas@gmail.com, [2]almeida@cdsid.org.br, [3]jonatasaa@yahoo.com.br
web-page: www.cdsid.org.br

ABSTRACT

The current demands of the business market reveal the relentless pursuit of organizations by coordinated changes in their strategic planning. In the electric sector environment, the problem of resource allocation is frequently faced by enterprises since there are several projects that disputes limited resources and it is so complex to search for the optimal distribution. Since these companies seek to minimize the operational cost and guarantee their services on market, their management policies meet different and often conflicting objectives. So, this work uses a decision model that allows a multidimensional evaluation in portfolio problematic with c-optimal portfolio concept in PROMETHEE V method. Then, this paper presents a new web-based DSS to apply this model, in which subjective criteria as well as financial, human and operational constraints might be considered in the problem formulation. The DSS uses graphic visualizations so that the decision aiding process is improved. The DSS is currently available on-line, by request. The use of the DSS and its model is shown by means of an application in an electricity company and their results are properly analyzed and show its real contribution.

Keywords: resource allocation, c-optimal portfolio concept, PROMETHEE, electricity company, web-based DSS, portfolio selection.

INTRODUCTION

Several moments in history have shown that electric energy has become an essential input to society so its contributions improve the productive sectors and, consequently, promote the socioeconomic development of nations. Currently, the electricity sector service covers a vast territorial and ever expanding so it is a key factor for all societies.

Although there are many natural resources to be exploited, the companies seek to improve its operational planning for generating energy. It should also be noted that political and economic factors make it more difficult for these organizations to ensure, simultaneously, lower operational costs and better services. In fact, the required resources are limited as well as consumer demand increases. For that reason is more complex to achieve an optimal distribution by selecting projects in the resource allocation context.

Due to these factors, the project selection can be analyzed in a multidimensional way and multicriteria decision making/aiding (MCDM/A) methods had been developed to assist this problem. These methods explores in many papers the application and benefits of MCDM/A principles and properties in order to support the decision making through a portfolio selection in energy companies [1]–[4].

This work is based on PROMETHEE V (Preference Ranking Organization Method for Enrichment Evaluation) method [5], whose approach uses non-compensatory rationality, using the c-optimal concept to overcome its scaling issues. Thus, this paper presents a web-based Decision Support System (DSS) to aid the decision maker (DM) in its process. It was designed so that the DM has its decision consolidated through the variety of tools offered by the DSS. The model was applied in an electricity company, which results and sensitivity analysis report are shown and discussed. Finally, the last section makes some final remarks, open questions and draws conclusions.

RESOURCE ALLOCATION IN THE ELECTRIC SECTOR: A LITERATURE REVIEW

Multicriteria approach to select a portfolio

In all decision making processes, DM's seek to prioritize their objectives and also achieve success of its organization by selecting projects. To do so, companies need to apply their resources, but the required amounts exceed the limits [6]. In spite of many procedures focused on resource allocation problems, it should be noted that procedures for portfolio selection in a multicriteria approach have been widely used in this context [7].
Thus, Multicriteria Decision Making is an approach based on decision maker's preference structure and involves value judgments of multiple objectives – which often conflict with each other – that need to be dealt with simultaneously [8].
Particularly, applications in the electrical sector were found in the literature and they were based on different methods. They differ from each other in how they estimate DM's preference and also aggregate their information to choose the best mix of projects that will be included in a portfolio. Among them, we can see works applying methods such as the Analytic Network Process (ANP) [1], Analytic Hierarchy Process (AHP) [2], additive method [3] and PROMETHEE V method [4].
In this work, PROMETHEE V deserves attention, because its applicability in portfolio problems is simple and easily understood by the actors of the decision process. In addition, criteria weights mean how each one is important to achieve DM's goals.
Its methodology is divided into two steps: the first one is based on pairwise comparison between projects (a_i) to measure outranking relations through net flows parameters ($\varphi_{(a_i)}$) according to l criteria of the problem. Then, on the second step, these parameters are maximized in an integer linear programming (ILP) so that the selected portfolio meets all the j problem's constraints [5]. The decision variable x_i indicates if a project a_i will be rejected (0) or inserted (1) into the portfolio (Equation 1).

$$\max \sum_{i=1}^{n} \varphi'_{(a_i)}\, x_i \quad - \quad .a. \quad - \quad \sum_{i=1}^{n} r_{ji} \cdot \quad B_j \quad - \qquad (1)$$

Mavrotas, Diakoulaki & Caloghirou [9] analyzed that formulation and found that a project a_i could never be selected if its net flow $\varphi_{(a_i)}$ was non-positive. This led them to propose a scale transformation to correct it by adding an increment $\left|\min \varphi_{(a_i)}\right| + \delta$ to the new scale. They noted that all projects could be selected in the ILP step of the PROMETHEE V method due to this transformation.
However, De Almeida and Vetschera [10] pointed out that the δ value on scale transformation can recommend different portfolios on the ILP step of PROMETHEE V. The portfolio values are impacted by the scale and they depend on the number of selected projects of the portfolio. In other words, this scaling issue can still provide distortions in the final result.

The c-optimal concept to solve scaling issues in PROMETHEE V method

In order to overcome the portfolio size effect, Vetschera and de Almeida [7] proposed the 'c-optimal portfolio' concept. They introduce this concept by inserting a new constraint on the second step of PROMETHEE V, so that a portfolio with exactly c projects is selected (Equation 2).

$$\sum_{i=1}^{n} x_i \leq \text{c} \qquad (2)$$

In practice, the original PROMETHEE V result select p projects and it is denoted here as a p-optimal portfolio. Once the distortion is centered on the portfolio size, the new constraint in Equation 2 will be added to the formulation (Equation 1) and the ILP step will run again to select an optimal portfolio with c projects. This value gradually increasing (c = p + 1, c = p + 2, c = p + 3, ...) as long as there is a solution.
It should be noted that a c-optimal portfolio remains unchanged for any value of δ. So, this concept overcomes the current distortion. Finally, the DM has to compare among all the solutions the best according to your goals.
In this sense, the proposed model for allocation of resources with the concept of c-optimal portfolio is divided in the following steps:

1. Calculate p-optimal portfolio (Equation 1);
2. Calculate viable c-optimal portfolios by inserting a new constraint (Equation 2) in the original formulation;
3. Calculate performance portfolio values P_k^C for each criterion k, where $v_{i,k}$ is the performance of a_i in criterion k;

$$P_k^C = \sum_{i=1;k=1}^{i=n;k=l} v_{i,k} x_i \qquad (3)$$

4. Calculate concordance indexes: measurement of how a portfolio is better than the other, according to the non-compensatory racionality;

$$C_{(c,p)} = \sum_{k=1}^{l} p_k \, , where \; P_k^C > P_k^P \qquad (4)$$

5. Recommendation: compare step 4 results to choose the best portfolio.

In order to help DM in the process in the c-optimal concept application, a web-based DSS has been developed incorporating graphic results and comparison in order to recommend the best portfolio. In addition, a sensitivity analysis may be done in the DSS.

A MCDM/A PORTFOLIO SELECTION DSS

Developed as a tool for applying the model with c-optimal concept, a Decision Support system (DSS) called "MCDM/A Portfolio Analysis with c-optimal portfolio concept for PROMETHEE web-based, code PU-PXCNG-WT1" considers the cognitive style of a user, by means of the way in which their observe and analyze information provided. Thus, the DSS meet several functional aspects, such as simplicity (it is friendly to the user), familiarity (its interface is comfort to the user) and flexibility (for navigation and input/output system).
In the input step, the required model entries are based on the necessary parameters to run this model, which includes information about the criteria, alternatives and constraints. The software input data is imported via excel spreadsheet in a single tab. Once these parameters are imported, the user can see a summary of the MCDM/A problem that will be evaluated by the model.

Then, the processing stage explore the preference relationships between the alternatives and, by the ILP step, the set of c-optimal portfolios is presented (see figure 1). In addition, the user can analyze thought a graphic analysis what is the best portfolio that achieve its goals.
Also, the DSS allows a sensitivity analysis, whose weights and performances can be varied by percentage. Graphic tools help this analysis and reinforce the portfolio recommendation, as seen in figure 1. All outputs can be exported as a report to the user.

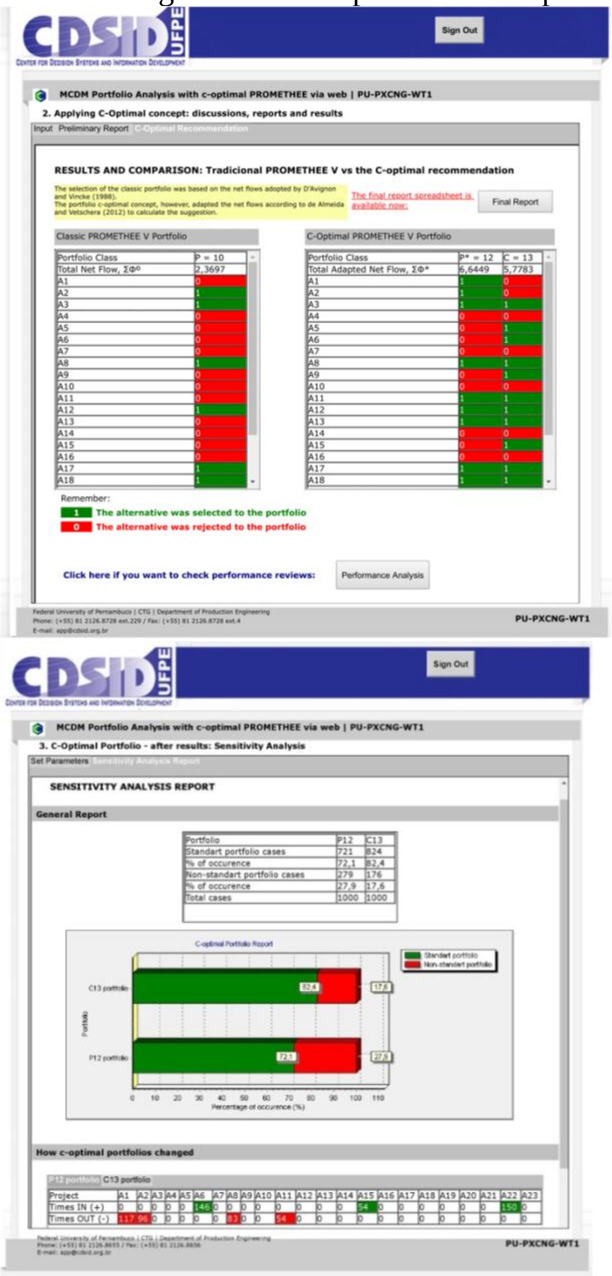


Figure 11: DSS screens of Portfolio Results Comparison with graph analysis and Sensitivity Analysis Report.

Developed as a web-based application, the DSS is available by request (see at www.cdsid.org.br).

STUDY CASE: APPLICATION IN AN ELECTRICITY COMPANY [4]

About the problem

In the context of an electricity company, which offer multiple services for consumers, this application is an adaptation from López and De Almeida [4] in their annual strategic review, in which about 83 new strategic projects need to be selected in a multidimensional environment that considers relevance and potential factors to many businesses of the group, such as: Impact on results; Strategic alignment; Improvement of regulated indicators; Contribution to customer satisfaction; Likelihood of achieving benefits; and Complexity. There are human and financial limitations and all them will be modelled in the problem as constraints.

Therefore, this work seeks to structure a multicriteria decision-making problem that allows the group to choose the best portfolio of projects. By implementing this portfolio, the organization is expected to materialize strategic planning in order to improve existing resources and enable such a sustainable growth.

Results and discussion

The numerical application thought the DSS presented three optimal portfolios: P37 (p-optimal), C38 and C39 (both c-optimal). The steps 3 and 4 of the model are so important to compare among the optimal portfolios what is better.

Through a performance comparison, by criterion, the concordance indexes show that almost 69% of the weights suggest C39 portfolio is better than C38. In addition, C39 is better than P37 in 86% of the weights, so the recommendation should be, in fact, the implementation of the strategic portfolio C39.

Among the possible solutions, the financial resources allocated to execute C39 portfolio are the smallest, while it requires low amount of human resources to be allocated. In any case, we must clarify that it is not the objective of the proposed model to obtain a solution that simultaneously presents better performances and lower allocable resources. Their concern aims to obey the problem constraints.

Then, it could be inferred that there is a society positive impact in the energy sector, once this study was applied in an electricity company. The implications of these results for practice are that when an adequate scale transformation is considered with c-optimal portfolio analysis, this can contribute to a company's permanent strategies of increasing productivity, as long as they maintain great quality indicators and also provide opportunities for sustainable growth.

Sensitivity Analysis results

This analysis was done in the DSS default mode simulating 100,000 instances of the c-optimal PROMETHEE model. All the results are arranged as shown in Table 1. Only the criteria weights varied within a range of ± 10% in a linear distribution of probability.

It should be noted that about 73% of all cases resulted in a recommendation of the C39 portfolio to the problem, while more than 80% indicated the same solutions as p-optimal and c-optimal solutions (P37 and C38).

Table 1 – Sensitivity analysis report from the application

Portfolio		P (P37)	C=P+1 (C38)	C=P+2 (C39)
standart portfolio	Frequency	88,214	85,396	73,015

(different class/ size)	Proportion	88.21%	85.40%	73.01%
non standart portfolio (equal class/ size)	Frequency	11,786	14,604	26,985
	Proportion	11.79%	14.60%	26.99%
Total		100,000	100,000	100,000

In this sense, that problem presents a strong robustness regarding the weights variation to solve the decision problem, since the impact observed with this variation did not generate great changes in the portfolio selection.

CONCLUSIONS

This work presents a new DSS applying the c-optimal portfolio concept for PROMETHEE method that follows all the described steps rigorously and overcome the portfolio size effects. Also, sensitivity analysis is introduced as a tool to improve the decision aiding process. Furthermore, graphic visualizations are introduced so that the entire layout are simple and rich in detail of the results to aid the decision maker by choosing the best portfolio. Graphic visualization has been increasingly applied as an important tool for the improvement of the decision aiding process, as shown by Roselli et al [11]. In this way, the DSS give safety and robustness to the recommendation, besides it provides users with clear and concise decision-making.
The application in an electricity company shows that the ideal c-concept for multidimensional problems is useful. Since, at worst, a solution of being applied is the original, so this leads one to believe that it does not make sense to apply PROMETHEE V without the c-optimal portfolio concept being actually associated.
Finally, the DSS presents a potential of benefits when applying multicriteria models for allocation of resources in many contexts, not only in the electrical one.

ACKNOWLEDGEMENT

This paper is part of a research study funded by the National Council for Scientific and Technological Development – CNPq and the Brazilian Council for Improving Higher Education – CAPES.

REFERENCES

[1] M. García-Melón, R. Poveda-Bautista, J. L. Del Valle M., “Using the strategic relative alignment index for the selection of portfolio projects application to a public Venezuelan Power Corporation,” *Int. J. Prod. Econ.*, vol. 170, pp. 54–66, 2015.
[2] P. Aragonés-Beltrán, F. Chaparro-González, J. P. Pastor-Ferrando, A. Pla-Rubio, “An AHP (Analytic Hierarchy Process)/ANP (Analytic Network Process)-based multi-criteria decision approach for the selection of solar-thermal power plant investment projects,” *Energy*, vol. 66, pp. 222–238, 2014.
[3] C. L. Martins, H. M. L. López, A. T. De Almeida, J. A. Almeida, M. B. De Oliveira Bortoluzzi, “An MCDM project portfolio web-based DSS for sustainable strategic decision making in an electricity company,” *Ind. Manag. Data Syst.*, vol. 117, no. 7, 2017.
[4] H. M. L. López, A. T. de Almeida, “Utilizando PROMETHEE V para seleção de portfólio de projetos de uma empresa de energia elétrica,” *Production*, vol. 24, no. 3, pp. 559–571, 2014.
[5] J. P. Brans, B. Mareschal, “Promethee V: Mcdm Problems With Segmentation

Constraints," *INFOR Inf. Syst. Oper. Res.*, vol. 30, no. 2, pp. 85–96, 1992.
[6] C. F. Larson, E. W. Gray, *Project managment: the managerial process*, 7ª. New York: McGraw-Hill Education, 2018.
[7] R. Vetschera, A. T. De Almeida, "A PROMETHEE-based approach to portfolio selection problems," *Comput. Oper. Res.*, vol. 39, no. 5, pp. 1010–1020, 2012.
[8] A. T. de Almeida, C. A. V. Cavalcante, M. H. Alencar, R. J. P. Ferreira, A. T. de Almeida-Filho, T. V. Garcez, *Multicriteria and Multiobjective Models for Risk, Reliability and Maintenance Decision Analysis*, 1ª., vol. 231. Cham: Springer International Publishing, 2015.
[9] G. Mavrotas, D. Diakoulaki, Y. Caloghirou, "Project prioritization under policy restrictions. A combination of MCDA with 0-1 programming," *Eur. J. Oper. Res.*, vol. 171, no. 1, pp. 296–308, 2006.
[10] A. T. De Almeida, R. Vetschera, "A note on scale transformations in the PROMETHEE v method," *Eur. J. Oper. Res.*, vol. 219, no. 1, pp. 198–200, 2012.
[11] L. R. P. Roselli, A. T. de Almeida, E. A. Frej, "Decision neuroscience for improving data visualization of decision support in the FITradeoff method", *Oper. Res. Int. J.*, 2019.

Multicriteria Web-Based Dss for Resource Allocation in Higher Education Organizations

Carolina Lino Martins [1], [2], Pascale Zaraté [1]

Adiel Teixeira de Almeida [2], Danielle Costa Morais [2]

[1]IRIT – Toulouse Capitole University – 2 rue du Doyen Gabriel Marty 31042 Toulouse Cedex 9 – France
Pascale.Zarate@irit.fr
web-page: https:// www.irit.fr/

[2]UFPE - Production Engineering Department, Federal University of Pernambuco
Av. Avenida da Arquitetura, s/n, Cidade Universitária, Recife-PE, Brazil
carol_tcch@hotmail.com, almeida@cdsid.org.br, daniellemorais@yahoo.com.br
web-page: http://www.cdsid.org.br/

ABSTRACT

The allocation of scarce resources is a complex problem, specially when it comes to budget constraints. Therefore, this work aims to propose a multicriteria web-based Decision Support System for resource allocation in the context of higher education organizations, more precisely, public universities that have budget constraints, such as Brazilian federal universities. To do so, a Brazilian federal university was chosen as a parameter to make a numerical application to validate the multicriteria model for resource allocation proposed and, afterward, a web-based DSS was developed. For the MCDM resource allocation model, an additive value function was considered to set the percentage of the total budget that every alternative should receive. The problem was seen as a particular case of project portfolio selection problem because its approach is deemed to be appropriate for a resource allocation decision context. The results were satisfactory, and the system provided a clear vision on how the resource allocation procedure works, the entire process became more transparent to the ones that are affected by it, to the decision makers and the government, enabling them to take more efficient and reasonable decisions.

Keywords: Resource allocation, Budgeting, MCDM /A, Universities, Web-based DSS, Model-driven DSS.

INTRODUCTION

Resource allocation problems usually involve conflicting decisions and are confronted by organizations of every size, type, and purpose, once their capacity to borrow funds or raise equity capital, for instance, has practical limits [1]. For public universities, it is even more challenging to deal with budgeting problems because they use their taxpayers' money to provide education services [2]. Hence, there is a tremendous societal interest (or at least should exist) in the way such money is allocated, where the cost of failure is seen as something unacceptable [3].

Multiple criteria decision making/aid (MCDM / A) area has been claimed as an effective way to assist decision makers (DMs) to deal with the challenges that involve resource allocation problems or budgeting problems [4], specially in higher education organizations [5]. Also, Turban *et. al*. and Power [6, 7] affirm that Decision Support Systems (DSS) can improve decision quality. Therefore, the use of a suitable web-based Decision Support System (DSS) meant to integrate MCDA / M analysis into the decision aiding process is an essential tool to respond to this ongoing challenge.
Multiple Criteria Decision Support Systems (MCDSS) are considered a "particular" type of system within the broad family of DSS [8]. MCDSS use different multicriteria decision methods to estimate efficient solutions, and they incorporate user's input in numerous phases of modelling and solving a problem [8].
According to [4], despite the growing attention to MCDA-based modelling approaches for resource allocation [9, 1, 10], there is still little indication in the operational research and decision sciences literature on how to structure these models in practice.
Besides, when considering the case of a university, the use of a suitable multi-attribute decision method integrated with a web-based Decision Support System to better distribute the limited budget could mean to reach the best compromise solution, to apply all the available resources with efficiency.
Thus, this work aims to fill this gap by proposing a multicriteria web-based Decision Support System to solve a resource allocation problem in the context of higher education organizations, more specifically, public universities that have budget constraints. For this reason, a Brazilian public university was chosen to validate the model; once, currently, the university does not have a model or system for such a problem. The research can contribute to the decision question of how to allocate universities internal budget properly [2].
The case study conducted in this work considers the results of previous studies related to MCDM portfolio selection problems, such as: [2, 11, 12, 13].

METHODOLOGY

The problem considered here is a special case of portfolio problematic, seen as a resource allocation problem. The decision maker of the problem was the director of the budget and planning department of the university, and the analyst was the author of this study. The university studied has 21 sectoral administrative units (called *UAS*) that are divided by areas, such as human sciences, biological sciences, engineering, faculty of medicine, etc. These administrative units are the alternatives, projects or budgetary units of the MCDM model.
Therefore, the set of alternatives is A = {UAS 1, UAS 2, UAS 3, UAS 4, UAS 5, UAS 6, UAS 7, UAS 8, UAS 9, UAS 10, UAS 11, UAS 12, UAS 13, UAS 14, UAS 15, UAS 16, UAS 17, UAS 18, UAS 19, UAS 20, UAS 21}.
The criteria of the model, defined by the DM, are evidenced in Figure 1.

Criteria	Description	Objective
InAlEqv	General index of equivalent students. Number of students entering, enrolled and graduated from undergraduate, postgraduate courses (master's and doctorate), and medical residences.	Maximize
IQCD	Faculty qualification criterion, that measures the academic staff qualification by the number of lecturers with Phd and master's degrees.	Maximize
IVO	Dropout rate criterion, defined by the summation of vacancies not filled in the regular admission process, plus vacancies arising from withdrawing, dismissed students and transfer.	Minimize
IPP	Total of research projects with external financial support.	Maximize
IPE	Total of extension projects with external financial support.	Maximize
ITS	Graduation success rate. Performance criterion indicator that measures the relationship between the number of graduates and the number of new entrants.	Maximize
IDEAE	Teaching efficiency. Measured by the relation between the total of equivalent students and the total of equivalent professors.	Maximize
IDGQ	Quality of the undergraduate courses, based on the evaluations from the National Institute of Studies and Educational Research Anisio Teixeira – INEP / Brazil.	Maximize
IDQM	Quality of the master's degrees courses, based on the evaluations from the Coordination for the Improvement of Higher Education Personnel – CAPES / Brazil.	Maximize
IDQD	Quality of the doctorate degrees courses, based on the evaluations from the Coordination for the Improvement of Higher Education Personnel – CAPES / Brazil.	Maximize

Figure 4: Criteria of the Model

The weights (scale constants) of the model were also defined by the DM and are presented in Figure 2.

Criteria	InAlEqv	IQCD	IVO	IPP	IPE	ITS	IDEAE	IDGQ	IDQM	IDQD
Weights	0,2505	0,0405	0,145	0,124	0,124	0,150	0,038	0,044	0,042	0,042

Figure 2: Weights of the Model

The model adopted was an additive aggregation procedure for portfolio problematic with compensatory rationality, because of the characteristics of the problem. The primary goal of the model is to maximize the objective function, considering the given constraints [1], that is a budget constraint. Therefore, the objective function (1.0) and the constraints (2.0) are written as:

$$\sum_{i=1}^{n} z_i\, v(A_i) \tag{1.0}$$

Subject to:

$$\sum_{i=1}^{n} z_i\, c_i \leq C \tag{2.0}$$

Where i represents every UAS from the university, z_i is defined as a binary variable indicating whether item A_i is included or not in the portfolio, thus $z_i = 1$ if it is included and $z_i = 0$ if it is not [14]. $v(A_i)$ is the value of item A_i obtained from the multi-attribute evaluation [15].

C and c_i are related to the constraints, where C is the budgeted amount available to fund all the UAS and c_i is the budget of each administrative unit, and it could be seen, for instance, as the cost to develop project i.

When considering a public university, no administrative unit can stay without receiving a part of the budget because of the minimum amount required to maintain the UAS, in services such as security, for example. Consequently, the decision problem here lies in defining which are the administrative units that will receive a part of the budget above the minimum value that each one must receive, that is, the total budget requested by the UAS, considering their performance for the set of criteria, and that is a project portfolio selection problem.

Moreover, to adequate the model in this study and taking into account equation (1.0) and inequation (2.0), the variables of the model can also be described as:

c_i = the budget requested by the administrative unit or the budget above the minimum limit that each UAS want to receive;

$min\ c_i$ = minimum percentage of the budget that each UAS should receive;

z_i = binary variable that is equal to 1 when the UAS will receive the requested budget or equal to 0 otherwise;

$z_i\ c_i$ = the budget allocated to UAS "i", which is equal to c_i when z_i is equal to 1;

B = total budget from the university available to be allocated;

C = total budget amount that is above the minimum percentage of the budget that each UAS should receive, that is:

$$B - \sum_{i=1}^{n} min\ c_i = C \qquad (3.0)$$

Finally, the evaluation results from an additive value function it is of the form [15]:

$$v(A_i) = \sum_{j=1}^{m} k_j v_j(x_{ij}) \qquad (4.0)$$

Where x_{ij} is the outcome obtained by item A_i in attribute j, v_j is the marginal value function of attribute j, k_j is the weight (scaling constant) for attribute j, and its summation must be equal to 1; $v(A_i)$ is the value of item A_i obtained from the multi-attribute evaluation [15].

The total budget available (B) considered for the problem was R$ 850,000.00, a value that represents 85% from 2018 total budget of the university, once that was the amount released by the Ministry of Education in 2018, due to government budget cuts. The minimum value considered that each administrative unit must receive ($min\ c_i$) was 70% from the last budget, a total of R$ 700,000.00 since that is the minimum amount considered to maintain the UAS. Thus, C = R$ 150,000.00. Following, the results from the model are shown in Table 1.

RESULTS

Table 1 - Resource allocation model results

Alternatives	$V_i(A_i)$	*Go?*	P_i %
UAS1	0,6682	0	12,16%
UAS2	0,5981	0	9,98%
UAS3	0,3880	1	2,92%
UAS4	0,3821	1	3,71%
UAS5	0,3804	1	3,58%
UAS6	0,2523	1	1,41%
UAS7	0,3422	1	2,55%
UAS8	0,2352	1	1,62%
UAS9	0,2799	1	1,75%
UAS10	0,2979	1	1,72%
UAS11	0,2319	1	1,57%
UAS12	0,4583	0	5,86%
UAS13	0,4044	0	4,72%
UAS14	0,4584	1	4,36%
UAS15	0,5781	0	10,43%
UAS16	0,7950	0	11,75%
UAS17	0,6195	1	9,78%
UAS18	0,3744	1	2,72%
UAS19	0,4103	1	2,58%
UAS20	0,2689	1	2,20%
UAS21	0,3339	1	2,62%
Total Value	5,2551		

Table 1 shows the alternatives, their respective value from the additive model (V_i) using a ratio scale normalization procedure, considered as the appropriate procedure for this type of problem [11], the percentage (P_i %) of the total budget that each administrative unit should receive, the alternatives that are selected to receive the budget amount above the minimum budget (*Go*?) and the total value, that represents the objective function of the model.

From the use of the additive model, the results indicate a portfolio with 15 projects, and in terms of budget value, the solution consumes R$ 148,831.30 from the available budget (R$ 150,000.00). When deeply analyzing the results, although UAS 1, for instance, has a satisfactory performance in the criteria considered, the summation of the results of more than one UAS brings additional outcomes to the overall portfolio value than choosing only one project with a valuable performance. Such a solution means that the resources are distributed in a more balanced way among all items of the portfolio, and that is the best compromise solution for the case.

Given this consideration on the problem, it was possible to implement the multicriteria web-based DSS, based on the user's primary needs. The name defined for the web system was: MDSSFRA (Multicriteria Decision Support System for Resource Allocation). The principal DSS component is the multicriteria model, and the target users are the administrative staff from the budgeting unit of the university considered, DMs participants from every UAS, since they are affected by the allocation procedure, facilitators, developers, and administrators.

To implement the multicriteria web-based DSS, a PHP web platform was developed on the server side integrated with Python and a Database system MySQL was applied to store and retrieve data using Structured Query Language (SQL). The components of the system are shown in Figure 3, and they evidence results from previous studies [2, 12, 13], which served as the basis for establishing Pages 1, 2 and 3 of the system, already explained in these previous studies. Page 4 was the result of the implementation of the multicriteria model developed by this study

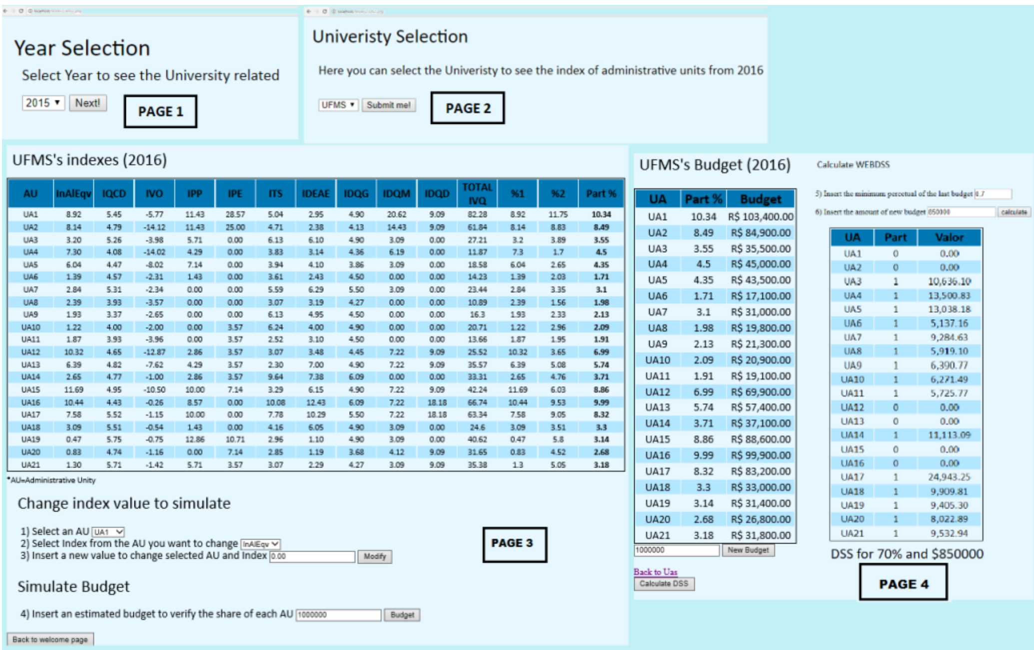


Figure 3 – Web system pages – user interface

Thus, Page 4 has two main tables. The first one shows the budget in financial and percentage terms and the possibility of simulating the results with a different budget. The last column (Budget) of the first table is the multiplication of the participation percentage of each UAS with the total budget available. The second table represents the MCDM model results evidencing the units that will receive a part of the budget above the minimum established by the university.

PERSPECTIVES AND CONCLUSION

This study presented a multicriteria web-based Decision Support System for resource allocation in the context of higher education organizations, more specifically, public universities that have budget constraints, with the aim of demonstrating how the use of a suitable multi-attribute decision method combined with a DSS could improve the distribution of a limited budget, which it could mean to reach the best compromise solution, by applying all the available resources with efficiency.

To validate the multicriteria model, a Brazilian federal university was chosen as a parameter to make a numerical application. The model was able to define the percentage of the budget that every budgetary unity of the university should receive. The model established 21 alternatives, which were the sectoral administrative units from the university, and ten criteria were defined by the DM, which was the director of the budget and planning department. A portfolio of 15 projects was found. Also, a multicriteria web-based DSS prototype was established for the problem.

The method proved to be valuable for managing the allocation of resources through a set of alternatives which were distributed rationally by explicit consideration of the real importance of the different criteria. An advantage provided by the system is that when there is a clear vision of how the resource allocation procedure works, the entire process becomes more transparent to the ones that are affected by it. Besides, the multicriteria web-based DSS could be used to provide background for the university considered or other universities when defining strategic resource allocation planning.

ACKNOWLEDGEMENT

This work is part of a research program funded by the Brazilian Research Council (CNPq), FACEPE (Foundation for support of Research by the State of Pernambuco) and Federal University of Mato Grosso do Sul (UFMS).

REFERENCES

[1] D. N. Kleinmuntz, Resource allocation decisions. In Edwards, W., Miles, R.F. and von Winterfeldt, D. (Eds), "Advances in Decision Analysis". Cambridge University Press, New York, NY, pp. 400-418, 2007.
[2] C. L. Martins, A. T. de Almeida, & D. C. Morais, Design of a Decision Support System for Resource Allocation in Brazil Public Universities. "International Journal of Decision Support System Technology (IJDSST)", 11(1), 20-34, 2019.
[3] G. Williams, Applying Management of Risk (M_o_R®) for Public Services. "GSW Consultancy", 2009.
[4] G. Montibeller *et al.*, Structuring resource allocation decisions: A framework for building multi-criteria portfolio models with area-grouped options. "European Journal of Operational Research", 199(3), 846-856, 2009.
A. Mustafa, M. Goh, Multi-criterion models for higher education administration. "Omega", v. 24, n. 2, p. 167-178, 1996.
[5] E. Turban, R. Sharda, D. Delen, "Decision Support and Business Intelligence Systems". 9th Edition. New Jersey: Prentice Hall, 2011.
[6] D. J. Power, Computerized decision support case study research: concepts and suggestions. In "Real-World Decision Support Systems" (pp. 1-13). Springer International Publishing, 2016.

[7] P. Korhonen, A. Lewandowski, J. Wallenius, Multiple Criteria Decision Support; Proceedings of an International Workshop, Helsinki, Finland, August 7-11, 1989. "Lecture Notes in Economics and Mathematical Systems". Springer-Verlag, 1991.
[8] K. Golabi, C. W. Kirkwoo, A. Sicherman, Selecting a portfolio of solar energy projects using multiattribute preference theory. "Management Science", v. 27, n. 2, p. 174-189, 1981.
[9] L. D. Phillips, C. A. Bana e Costa, Transparent prioritization, budgeting and resource allocation with multi-criteria decision analysis and decision conferencing. "Annals of Operations Research", v. 154, n. 1, p. 51-68, 2007.
[10] C. L. Martins *et al.*, An MCDM project portfolio web-based DSS for sustainable strategic decision making in an electricity company. "Industrial Management & Data Systems", v. 117, p. 1362-1375, 2017.
[11] C. L. Martins *et al.*, Decision Support System for resource allocation in Brazil public universities (short paper). "Proceedings of the 2017 International Conference on Decision Support System Technology: Data, Information and Knowledge Visualization in Decision Making". Isabelle Linden, Shaofeng Liu, Jason Papathanasiou, Christian Colot (Eds.), ISBN: 978-2-917490-28-0, p. 119-124, may 2017.
[12] C. L. Martins, P. Zaraté, A. T. de Almeida, D. C. Morais, K. C. Peters, A web-based DSS for resource allocation in a Brazilian public University (short paper). "Proceedings of the 2018 International Conference on Decision Support System Technology: Sustainable Data-Driven & Evidence-based Decision Support with applications to the Environment and Energy sector". Panagiota Digkoglou, George Tsaples, Fatima Dargam, Isabelle Linden, Bertrand Mareschal, Jason Papathanasiou (Eds.), ISBN: 978-2-917490-29-7, p. 304-310, may 2018.
[13] R. T. Clemen, J. E. Smith, On the choice of baselines in multiattribute portfolio analysis: a cautionary note, "Decision Analysis", Vol. 6 No. 4, pp. 256–262, 2009.
[14] T. de Almeida, R. Vetschera, J. A. Almeida, Scaling issues in additive multicriteria portfolio analysis, in Dargam, F. et al. (Eds), Decision Support Systems III – Impact of Decision Support Systems for Global Environments, "Lecture Notes in Business Information Processing", Vol. 184, Springer, Cham, pp. 131-140, 2014.

An on-line platform for supporting DSS-research collaboration (EWG-DSS Collab-Net Version 5)

Jean Gomes Turet[1], Jadielson Moura[1], Ana Paula Cabral Seixas Costa[1], Fatima Dargam[2], Pascale Zaraté[2], Isabelle Linden[2]

[1]Universidade Federal de Pernambuco
Management Engineering Dep., Av. da Arquitetura, S/N,
50740-550, Recife-PE, Brazil [2] EURO Working Group on
Decision Support System/Coordinators jeanturet@gmail.com;
apcabral@cdsid.org.br
f.dargam@simtechnology.com, pascale.zarate@univ-tlse1.fr,
isabelle.linden@unamur.be

ABSTRACT

This paper reports the development progress of the project Collab-Net from the EWG- DSS in its current fifth version. The Collab-Net project aims at investigating the publication relationships in the DSS community in an automatic way, allowing researchers to analyze their own collaborative network. Moreover, in version 5, a web- based platform was developed and deployed in order to allow better possibilities of collaboration, in terms of joint-published work, among DSS researchers in their updated areas of work, in a portable and convenient way at any time. This paper presents the web- based platform development of the EWG-DSS Collaboration Network Project (EWG- DSS-Collab-Net), showing its new trends and advances (network collaboration, dashboard collaboration and word map), by taking in account the previous studies and versions of the Collab-Net project.

Keywords: Collab-Net, Network Collaboration, EWG-DSS

INTRODUCTION

Nowadays, it is important to identify the collaborative network between researchers. These networks provide important mechanisms of analysis since it will be possible to identify patterns as well as the connection between the researchers. Currently, it focuses on how the general structure affects researchers within research groups, scientific groups, among others [1] [4]. In this context, collaboration networks are tools that bring important mechanisms to improve academic productivity [2] [3].

Based on the importance of collaborative analyses, the EWG-DSS Collab-Net 5 was developed. The Collab-Net V.5 platform aims to investigate the publication relationships in an automatic way, allowing researchers to analyse their own collaborative network. Moreover, also allows possibilities for future collaboration among EWG-DSS members only using a Web-based platform, anywhere at any time.

This project has already advanced in development through four versions with various publications [9] [10] [11] [12] [12] [14] [15] [16], and is now in its fifth version, bringing greater possibilities for the DSS community members to check their collaborative network from online authoring platforms, like Google Scholar. The main feature of this new version is the possibility of visualizing the collaborative network from graphs and figures that bring higher levels of interaction among group members.

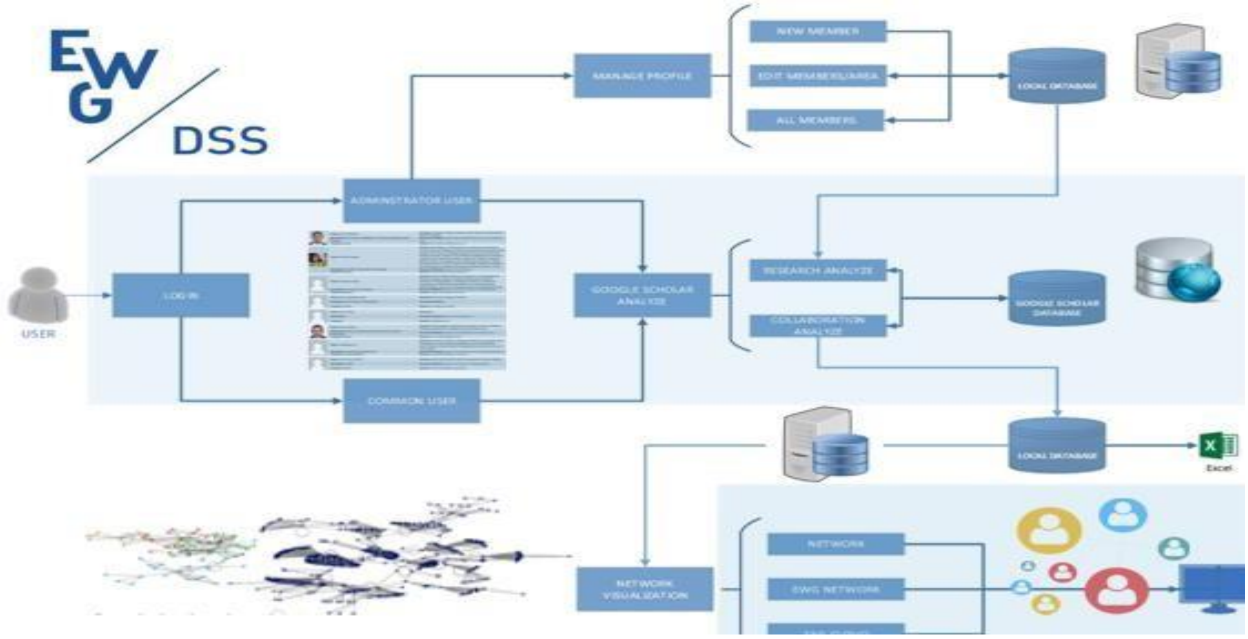
Thus, this paper describes the new recent web-based platform developed of the EWG- DSS Collaboration Network Project (EWG-DSS-Collab-Net), Collab-Net V.5 showing its new trends and advances if compared with the other versions.

COLLAB-NET V.5

Based on previous versions, Collab-Net version 5 was developed maintaining its original purpose of enabling the affiliated members to investigate the publication relationship of the collaborative interaction among papers authors, within a publication database from local database and google scholar. In addition, this version improves the process of data automation, increases the interaction between the members, brings significant changes in the design, allows the user to export the collected data to CVS Excel format and the user can see its graphical information about the nodes connections of own member or other member network (collaborative network).

In the Collab-Net version 5, the entity-relationship model is the same of the Version 3. The Information relating to authors, areas of knowledge, document published, keywords and logs of the system entities could be stored in their specific table with some relevant data from Google Scholar. Once developed the conceptual data model, all feature was implemented in the MySQL database using Structured Query Language (SQL).

Fig. 1. Web-Based Collab-Net System Flow



Once the database was ready, the web-based system can be started. Differently of the Collab-Net Version 3, the Collab-Net Version 5 was developed with different tools. First, the PyCharm IDE was used to develop the Web application, incorporating the Python development language together with Django Framework. The Bootstrap by Twitter was used to Collab-Net design. Therefore, all platforms of software development and database systemare free, which means that are not limited use or constraints. Despite it was the main motivation of these technology usages, other key motivation also be relevant, such as: easy integration these technologies, reliability and portability of Django Framework, further online support and established technologies.

As a way to present the navigability in the web-based Collab-Net system. Figure 1 shows the flow of both administrator and affiliated members' profiles when using the system. It shows from the member login until the possible results of the searches, as well as all functionalities of the system.

As a resulting from the software development process, a web-based system to support the needs of Collab-Net project was available to the affiliated members. This system offers a simple and integrated environment to allow affiliated researchers to analyse their own collaborative network anytime, anywhere in all platforms (Desktop, Laptop, Smartphones and Tablet), as well as possibilities for future collaboration among EWG-DSS members. The structure is presented in detail below.

COLLAB-NET VERSION 5 STRUCTURE

The Collab-net has a new login page. On the new login page, the user can access the Collab-Net V.5 environment from a previous registration (performed by the administrator). The user will automatically be directed to the administrative version of the system or to the normal version of the system.

After the user logs in, this is redirected to the Collab-Net v.5 home screen (Fig. 2). For admin users this part is more complete. There is a possibility of registering new users to the system, as well as the possibility and edition of new members and areas of action. The conventional user does not have these options; however, he can perform the analysis from google scholar, add new works in the database and check his collaboration network.

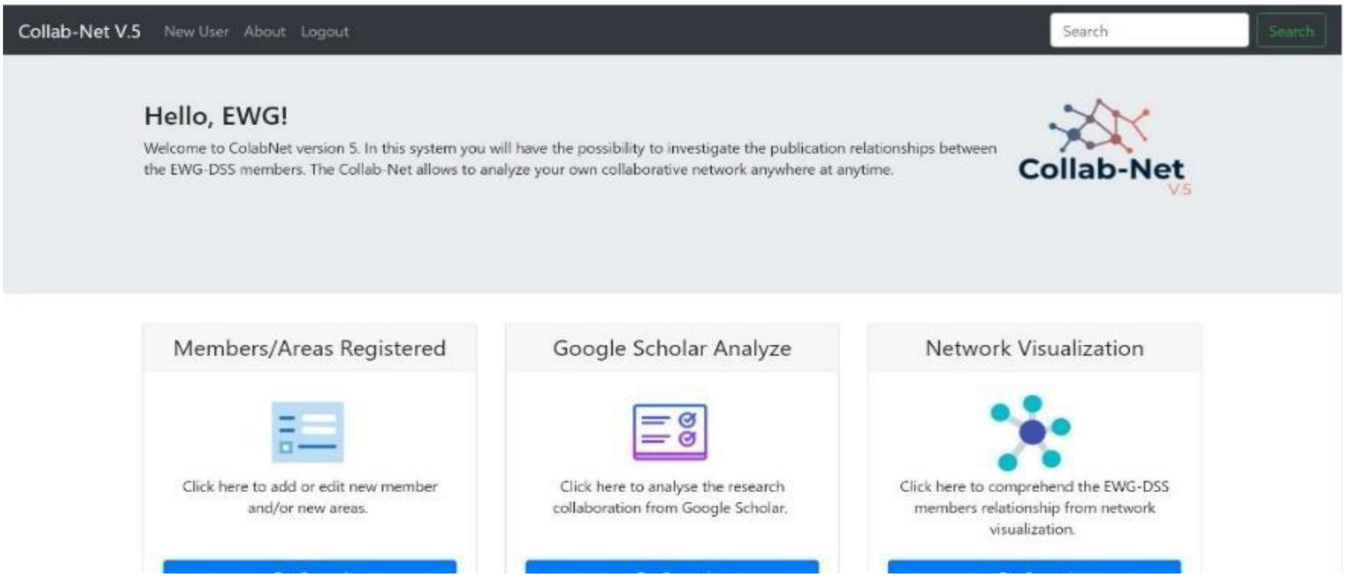


Fig. 2. Collab-Net V.5 Home Screen

One of the main parts of Collab-Net version 5 is in Google Scholar. In this part of the system the user can perform two types of queries: to analyse the registered members in the system, as well as verify the collaboration with other members.

First of all, the user can verify relevant information from any member that is registered in Collab-Net version 5. From the researcher's choice, information such as the researcher's full name, affiliation, citation numbers, e-mail, factor H will be displayed.

In this way, there is greater interaction among members since they may have greater knowledge of each other. In addition, all publications will be displayed, informing title and year (Fig. 3). These data are not saved in the local database but are used for consulting purposes only.

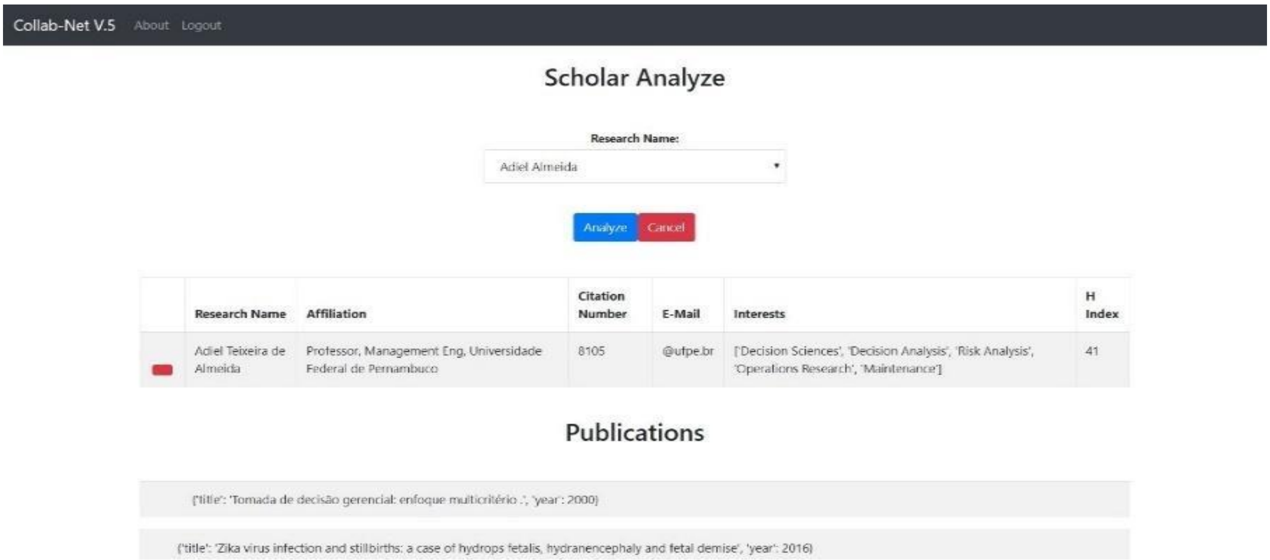


Fig. 3. Collab-Net V.5 Individual Scholar Analyse

In a second moment, the user can identify the collaboration with the other registered members. The system will connect to google scholar and conduct the query. The result will come in excel format and/or in a dashboard which will include the name of the main author, related publications and the co-authors involved. This data will also be saved in the local database to future analyse.

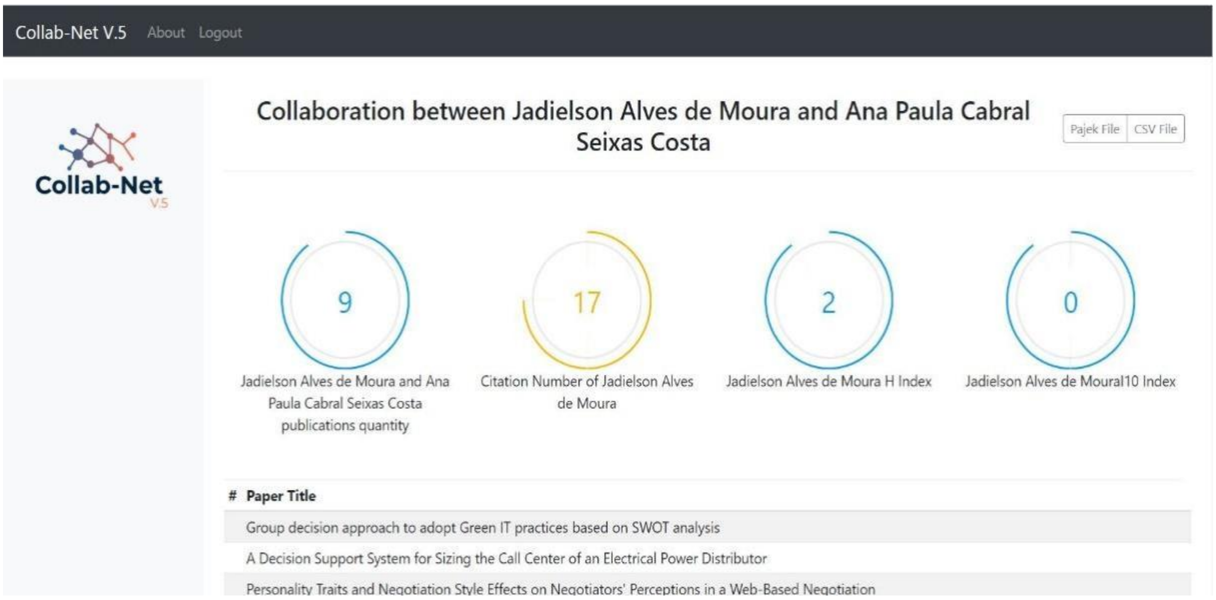


Fig. 4. Collab-Net V.5 Collaboration Analysis

The main part of this version is in collaborative networks. Collaborative Networks are defined as an intra or inter-organizational set with a common objective, obtaining collective solutions. In this way, the main benefits of Collaborative Networks are the sharing of knowledge (and learning by individuals), trust established between relationships and, finally, the enhancement of strategic skills of companies through learning [5][8].

In this context, Collab-net version 5, in addition to bringing the conventional form of collaboration analysis, allows members to identify their collaboration network from graphs and the collaboration network of the EWG Group. The members can have graphical information about the nodes connections of own member or other member network. For this, some fundamental Pajek[6][7] principles are used.

Pajek is a network analysis and visualization program specially designed to handle large data sets. Main objectives in the Pajek project are: 1) to facilitate the reduction of a large network into several smaller networks that can be further treated using more sophisticated methods 2) to provide the user with powerful visualization tools 3) to implement a variety of efficient network algorithms [6][7].

Based on these principles, Collab-Net V.5 refines the data that is saved in the database. This happens so that the result is faithful, since any change in the data, the collaboration network may be wrong. The system has a refinement algorithm for creating a pattern in member names. This algorithm is based on open source google refine, a google algorithm for data refinement.

This way, the user can generate his network quickly from three types of graphs at the discretion of the user (Fig. 5), starting from his publications and co-authors (only those who belong to the group and are registered in the system).

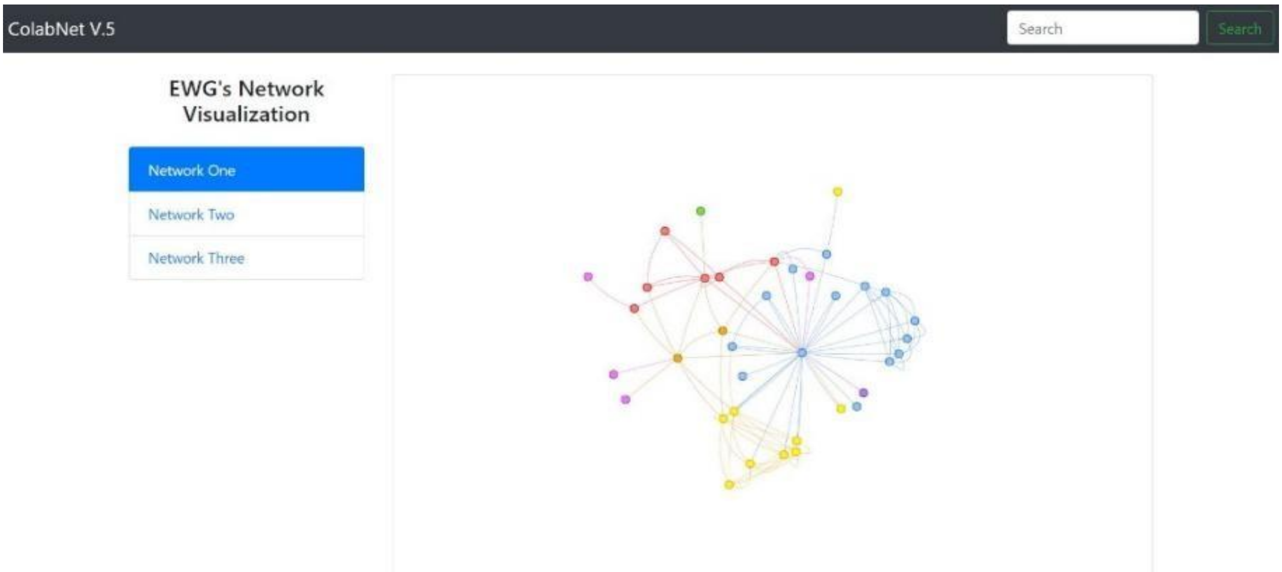


Fig. 5. Collab-Net V.5 Network Visualization

Moreover, the user can also analyse the EWG-DSS collaboration network in a general way. Although this part is still under development, specifications and initial coding were already provided for its proof of concept. There are constant improvements in the data refinement algorithm and in the collaborative network elaboration algorithm. The section 4 will present a preliminary study with this functionality.

CONCLUSIONS

In this paper, we described the new advances of the Collab-Net project in a current extended version with a web-based system. The Collab-Net 5 has deployed a more trendy environment with new tools, including a collaborative network that enables members to acquire more information about collaboration among DSS community members in a simple, effective and fast way. For the implementation of the current platform version, Python Language, Django Framework and Bootstrap by Twitter were used. Among other relevant and efficient solutions brought to the platform with these tools, they also have contributed to improve the Collab-Net V.5 with novel features, as the word map for instance. The word map brings new possibilities to visualize and analyse the main areas and topics of research, being more intensively investigated by the DSS community.

In a future work, new features still need to be implemented in the web-based platform. The focus of the Collab-Net future versions developments targets on: mechanisms to filter and transform the collected data; support the data collection process in more than one publication database; provide an intra-communication mechanism to encourage members to interacting via the available web-based platform and enable the sending of invitation to affiliation.

REFERENCES

[1] Sun, S., Manson, S.: Social network analysis of the academic GIScience community. The Professional Geographer 63(1), 18–33 (2011)

[2] Newman, M.E.J.: The structure and function of complex networks. SIAM Review 45(2), 167–256 (2003)

[3] Freeman, F.: Visualizing social networks. Journal of Social Structure 1(1) (2000)

[4] Shi, Q., Xu, B., Xu, X., Xiao, Y., Wang, W., Wang, H.: Diversity of social ties in scientific collaboration networks. Physica A 390, 4627–4635 (2011)

[5] Newman, M. E. Coauthorship networks and patterns of scientific collaboration. Proceedings of the National Academy of Sciences of the United States of America, v. 101, n. Suppl 1:5200-5205, april 2004.

[6] Nooy, W; Mrvar, A; Batagelj, V. Exploratory Social Network Analysis with Pajek. Cambridge University Press, 2011.

[7] Mrvar, A., Batagelj, V.: PAJEK - Program for Large Network Analysis (1996-2013), http://pajek.imfm.si/doku.php, http://vlado.fmf.uni-lj.si/pub/networks/pajek/

[8] Hansen, D., Shneiderman, B., Smith, M.: Analyzing Social Media Networks with NodeXL: Insights from a Connected World2. Morgan Kaufmann (2010) ISBN-10.

[9] Dargam, F., Ribeiro, R. and Zaraté, P., “A Collaboration Network for the EURO Working Group on DSS”, ISMICK 2008, International Symposium on the Management of Industrial and Corporate Knowledge, Niteroi, Rio de Janeiro, 3rd-5th Nov. 2008.

[10] Dargam, F., Ribeiro, R. and Zaraté, P., “Towards a collaboration network for the EURO Working Group on DSS. (Advances of the Project)”, Proceedings of the 23rd European Conference on Operational Research, EURO XXIII, Bonn, July 2009.

[11] Dargam, F., Ribeiro, R. and Zaraté, P., “How does the EURO Working Group on DSS interact? A social academic network analysis 1998-2008”, Proceedings of the 24th European Conference on Operational Research, EURO XXIV, Lisbon, July 2010.

[12] Bouaziz, R., T. Simas, T, Dargam, F., Ribeiro, R, Zaraté, P., “A Social-Academic Network Analysis of the EURO Working Group on DSS”. International Journal of Decision Support Systems Technologies, Vol.2, Nr. 4, pp. 13-36, IGI-Global, 2010.

[13] Dargam, F., Ribeiro, R. and Zaraté, P., “Networking the EWG-DSS: How do we proceed now?” (Short Paper) in (Eds.: F. Dargam; J. Hernández; S. Liu; R. Ribeiro; P. Zaraté) of the Proceedings of the EWG-DSS London-2011 Workshop on Decision

Systems , June, 2011, IRIT, Université Paul Sabatier, Report IRIT/RR--2011-14-FR; 2011.

[14] Dardenne, D., Dargam, F., Linden, I., Liu, S., Ribeiro, R., Sun, W., Zaraté, P., "Extending the Analysis of the EURO Working Group on DSS Research Collaboration Network (EWG-DSS Collab-Net V.2)" in Proceedings of the 25th European Conference on Operational Research, EURO XXV (Stream: Decision Support Systems), Vilnius - July 8th- 11th, 2012.

[15] Dargam, F., Ribeiro, R., Zarate, P, Liu, S., "A social network perspective of DSS research collaboration in Europe". IRIT, Université Paul Sabatier, Research Report IRIT/RR--2013-27-FR; 2013.

[16] F.C.C. Dargam, Isabelle Linden, Shaofeng Liu, Rita A. Ribeiro, Pascale Zaraté, "The Development Roadmap of the EWG-DSS Collab-Net Project: A Social Network Perspective of DSS Research Collaboration in Europe", in Springer Series LNBIP 164 - Lecture Notes in Business Information Processing (LNBIP_DSS) on "Decision Support Systems II - Recent Developments Applied to DSS Network Environments". ISBN: 978- 3-642-41076-5. DOI: 10.1007/978-3-642-41077-2. Eds.: Jorge E. Hernández, Shaofeng Liu, Boris Delibasic, Pascale Zaraté, Fatima Dargam, Rita Ribeiro. LNBIP 164. Springer Publisher, 2013.

Forecasting, Performance-based & Optimization Systems

A benchmark on a new approach for the automation of energy forecasting

Sabine Ritter, Martin Seidler
ifesca GmbH
Am Vogelherd 10, 98693 Ilmenau, Germany
sabine.ritter@ifesca.de, martin.seidler@ifesca.de
web-page: https://www.ifesca.com

Maria Jüttner
Leipzig, Germany
maria.juettner@gmx.de

ABSTRACT

Each solid planning process in the energy sector is based on forecasted energy data such as energy demands. These forecasts mainly rely on historical data exchanged between participants of the energy market.

In the underlying forecasting processes, the data must first be collected in order to be checked for errors and corrected if necessary. Currently, this data processing step can be easily automated. However, subsequent decisions required in these forecasting processes are typically performed manually by human operators. Such decisions include the analysis of characteristics specific to each energy data set, the selection of suitable forecasting models, the configuration and execution of said models, the monitoring of forecasting results, and the (daily) correction of forecasting model configurations. These decisions are compounded by the rapidly growing amount of energy data available for forecasting due to the widespread introduction of smart meters and the ubiquity of Internet-of-Things (IoT) Services. As a result, the demand for human resources for the performance of forecasting tasks is likewise rapidly increasing.

In this paper, a new intelligent decision support system (IDSS) for the automation of the entire forecasting process on energy data is presented. This new approach is compared to the manual forecasting process with respect to the amount of time required to generate and the quality of forecasting results. For comparison purposes, a fixed set of test data is used by both the new approach and a human operator (Maria Jüttner, forecasting expert at a municipal utility company in central Germany) for the generation of forecasting results.

Keywords: Forecasting, Energy Forecasting, Artificial Intelligence, Contest Human vs. Machine, Contest Natural vs. Artificial Intelligence, Decision Support System

HISTORICAL OVERVIEW: 30 YEARS OF FORECASTING

This chapter provides a small overview of forecast competitions and their influence to forecast research. The field of time series based forecasting gained more attention in the early 1980s and so the "International Journal of Forecasting" [1] and "Journal of Forecasting" [2] were founded. At this time, Spyros Makridakis announced the first M-Competition, an empirical study that has compared the performance of a large number of major time series methods. Until today, there were three more M-Competitions.

The M1-Competition [3], held in 1982, used a static data set of 1001 time series taken from demography, industry and economics. Forecasts were made with statistical and judgmental methods, for example the AutoRegressive-Moving Average (ARMA), various exponential smoothing and regression models. Especially the ARMA models were parameterized by hand and needed about one hour each. As stated from [4], the competition caused researchers to:

- focus attention on which models generated good forecasts, rather than on the mathematical properties of those models
- consider how to automate forecasting methods
- be aware of the dangers of over-fitting
- treat forecasting as a different problem from time series analysis

The M2 [5], held from 1987 to 1991, used only 29 time series but in a real-world scenario. The data was updated after one year to get new forecasts. Furthermore, the competitors were allowed to use additional information to improve the predictive accuracy.

The M3 [6] was intended to both replicate and extend the features of the M1-Competition and the M2-Competition. The data set increased to 3003 time series taken from micro, industry, macro, finance, demographics, and other domains. The extension involved the inclusion of more methods and researchers, in particular in the areas of neural networks and expert systems. It turned out that methods that are more sophisticated outperform the simpler ones. The best forecasting model was Theta [7] [8], followed by the commercial software package ForecastPro [9]. ForecastPro used a master controller to determine the model family and afterwards an automatic model identification based on the Bayesian Information Criterion. Thus, each model is individual.

The latest M4-Competition [10], held in 2018, utilized 100.000 real-life time series and incorporated all major forecasting methods, including those based on Artificial Intelligence (Machine Learning, ML), as well as traditional statistical ones. The dataset was mainly taken from economic, financial, demographical and industrial areas, while also including data from tourism, trade, labor and wage, real estate, transportation, natural resources and the environment. The most accurate methods were combinations of statistical approaches. However, the best model was a hybrid [11] that utilized statistical models with features learned by a recurrent neuronal network (RNN). Therefore, the forecast is created from global parameters (weights of the RNN) and per-time-series parameters like seasonality- and smoothing-coefficients (from statistical approach).

We are convinced that these hybrid and automated approaches can be enriched by more methods from the field of Artificial Intelligence.

INTRODUCTION

Forecasting especially of energy data is of particular importance in the energy sector as it provides the basis for every decision in all concerned processes such as generation management, energy sales, expansion planning, grid management, balancing group management, and energy procurement. Due to an immensely increasing amount of available energy data caused by widespread usage of smart meter and IoT-technologies, the time required by human operators for these forecasting issues is massively increasing as well. A new intelligent decision support system (IDSS) [12] is now offering a solution for this problem. Based on methods of Artificial Intelligence (AI), the IDSS reduces the complexity of forecasting processes for human operators to a minimum by maximizing the automation of the process: The IDSS is able to analyze and process input data fully automated, for instance, by identifying data gaps and generating adequate substitute values. Based on these analysis results the IDSS automatically performs the selection, the configuration and the training if needed of the best-fitting forecast method for every specific input data set.

This paper presents a comparison between the performance of this new IDSS and the performance of a human operator for a realistic forecasting scenario with real energy data. Therefore, the human operator Maria Jüttner, forecasting expert at a municipal utility company in Central Germany, as well as the IDSS performed the whole forecasting process for an energy data set including five time series and respectively two forecasting time horizons. The results are compared with respect to quality and time amount needed to generate them.

Contest Scenario

The data set in this paper consist of five time series containing energy data with a sample frequency of 15 minutes and can be seen in [13]. They were divided in a training set, which contains two times 11 months of data and three times 12 months of data and a testing set. These long measurements depict any behavior that can be observed for each series. The set reflects typical use cases in the energy market and is composed as follows:

ID 100093: Research institute; ID 100122: Authority; ID 100500: Logistics center;

ID 100750: Retirement home; ID 100755: Office building

The data was available in different formats because it was collected from various German costumers. Each series was divided into a training and a testing set. With this data set the contestants were asked to generate a short-term forecast (48 hours) and a mid-term forecast (1 month). Furthermore, no additional data was provided.

DISCUSSION

Following a short summary of the approach of the IDSS is given. As a first step, the structure of each input file is analyzed in order to load the correct data into the database. This step can be annoying due to different formats, DST-change and corrupt files. Afterwards a divide and conquer approach is used. Many statistical tests and transformations were done with the aim of data understanding, detrending and pattern recognition. In dependence of the results of the analysis and the desired time horizon of forecast, the forecast model is selected and parameterized. The hybrid model is build out of regression, random forests and artificial neuronal networks (ANN). Finally, the model is validated through hold out data. All tasks are computed highly parallel at cloud servers with Intel Skylake Xeon CPUs.

In contrast, the human operator Maria Jüttner is doing the same steps manually and sequentially using her experience to check the data of each time series and select a model. Her presented forecasts were made with random forests at a desktop PC powered by Intel Core i5. Both IDSS and the human operator use Python as modeling language.

An example for the forecasted time series with ID 100750 is shown in Figure 1, comparing the results of the human operator Maria Jüttner (MJ) and the IDSS over a period of two weeks out of the month-ahead horizon. The complete results of this experiment can be seen in [13].

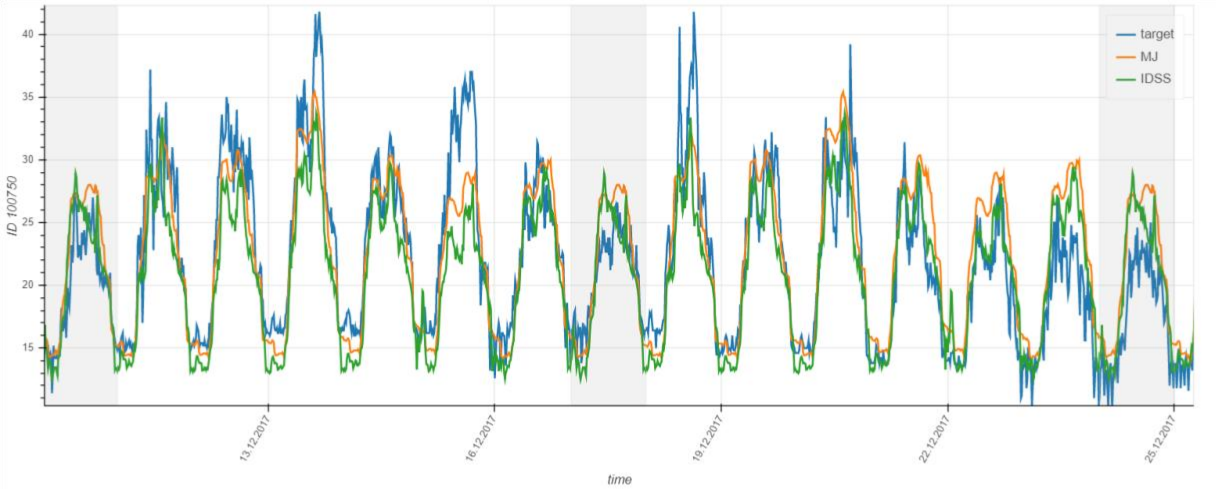


Figure 12: Detail of forecasting results for ID 100750 (month-ahead forecast)

Table 1 shows the results for the time required for the several steps within the forecast process.

Table 5: Execution Time for data sets in minutes for both time horizons

	ID 100093		ID 100122		ID 100500		ID 100750		ID 100755	
Process step	**MJ**	**IDSS**	**MJ**	**IDSS**	**MJ**	**IDSS**	**MJ**	**IDSS**	**MJ**	**IDSS**
Data input	**0,62**	0,74	0,60	**0,50**	**0,58**	0,68	0,53	**0,47**	0,67	**0,48**
Plausibility check	15,00		20,00		20,00		15,00		15,00	
Data correction	2,00		1,50		2,00		2,00		1,50	
Data analysis	30,00		35,00		30,00		30,00		35,00	
Forecasting method selection and configuration	0,92	**1,47**	0,87	**1,95**	0,62	**1,75**	1,07	**1,95**	0,94	**2,15**
Forecasting method training	0,60		2,02		1,34		0,62		0,87	
Forecasting execution	< **0,01**	0,17	< **0,01**	0,17	< **0,01**	0,15	< **0,01**	0,13	< **0,01**	0,13
Validation	5,00	**0,02**	5,00	**0,02**	5,00	**0,02**	5,00	**0,01**	5,00	**0,02**
Total	54,14	**2,40**	64,99	**2,63**	59,54	**2,60**	54,22	**2,57**	58,98	**2,78**

The first significant differences can be seen for the data preprocessing including data input, plausibility check, data correction, and data analysis. The step of data input may cause problems due to import libraries not being flexible and/or robust. In addition, the following steps of correcting the input data and analyzing the underlying characteristics are very time-consuming if executed manually. This information is especially important for the decisions to be made concerning the best forecasting model and its parametrization, though. The automation of these work steps can save an enormous amount of time as can be seen in Table 1. Furthermore, the parallelization of several processing steps as well as the scaling across random quantities of forecasting tasks via cloud architectures are additional advantages of the IDSS. Therefore, the IDSS has the ability to handle even forecasting of Big Data volumes. Conversely, the human processing can not be executed parallel and is scalable only to a very limited extent. The differences in the execution time of the forecasting process, which are influenced by the different computational technologies, are negligible in contrast. Thus, the different computational technologies can be neglected as well.

An example of the quality results can be seen in Table 2 for the day-ahead forecasting horizon and in Table 3 for month-ahead forecasting horizon.

Table 2: Benchmark for data set for day-ahead forecast (48 hours)

	ID 100093		**ID 100122**		**ID 100500**		**ID 100750**		**ID 100755**	
Error metric	**MJ**	**IDSS**	**MJ**	**IDSS**	**MJ**	**IDSS**	**MJ**	**IDSS**	**MJ**	**IDSS**
MEAN (Mean Error)	**-1,89**	6,78	**0,07**	0,82	1,33	**0,89**	1,67	**-0,02**	**0,06**	0,11
MAPE (Mean Average Procetual Error)	**2,72**	5,95	5,71	**3,57**	7,92	**5,82**	14,36	**11,32**	**7,74**	8,74
NRMSE (Normalized Root Mean Square Error)	**0,04**	0,73	0,08	**0,05**	0,1	**0,09**	0,18	**0,17**	**0,11**	0,13
APE (Absolute Procentual Error)	**2,76**	5,76	5,72	**3,53**	6,76	**5,41**	14,76	**12,19**	**7,53**	8

Table 3: Benchmark for data set for month-ahead forecast (4 weeks)

	ID 100093		**ID 100122**		**ID 100500**		**ID 100750**		**ID 100755**	
Error metric	**MJ**	**IDSS**	**MJ**	**IDSS**	**MJ**	**IDSS**	**MJ**	**IDSS**	**MJ**	**IDSS**
MEAN (Mean Error)	8,61	**-1,02**	0,64	**1,04**	1,83	**0,64**	0,82	**-0,74**	-0,35	**-0,27**
MAPE (Mean Average Procetual Error)	6	**3,8**	6,22	**4,94**	11,56	**9,41**	**11,75**	11,99	12,05	**11,45**
NRMSE (Normalized Root Mean Square Error)	0,07	**0,05**	0,09	**0,06**	0,15	**0,12**	0,46	**0,17**	0,23	**0,21**
APE (Absolute Procentual Error)	5,76	**3,83**	6,29	**4,87**	9,92	**8,11**	**11,87**	12,59	13,86	**12,65**

In general, the short-term results are better than the long-term results for both competitors. Moreover, the results reach a similar level due to forecasting precision. Thus, for comparison different error metrics have to be taken into consideration. The Mean Error (MAPE) is a good first indicator, but penalizes errors in the lower load range stronger. In contrast, the Normalized Root Mean Square Error (NRMSE) disregards the relative load status, which can be seen in the results [13]. In summary, the results of the human operator Maria Jüttner outperform the IDSS for two time series in the short-term sector. In the mid-term sector, the results of both competitors are almost similar for one time series. In the other 7 out of 10 cases, the forecasting results of the IDSS have a higher quality, almost independent of the considered error metric.

CONCLUSIONS AND FUTURE WORK

The data set in the scenario presented in this paper is quite typical for the use cases in the energy market: The time series are cyclical, just sparsely noisy, and contain a big amount of information well detectable and exploitable by machines. Furthermore, there is no additional data to be taken into account within the presented scenario. Therefore, the classical model approaches used by the IDSS are applicable for this data and the results of the IDSS are convincing: The automated forecasting process for the given input data is performed within just a few seconds, providing forecasting results that are approximately similar to the results of the human expert concerning quality.

Numerous data sets do not meet those criteria, though. These data sets need extensive investigations potentially with analytical methods usually not used for data analysis. In these cases, automated data processing may not be capable of successfully determining acceptable results. Here human experts may have advantages due to intuition and expert knowledge.

Therefore, further experimental studies with various input data are going to be performed including for instance different time horizons to be forecasted and additional data such as meteorological data or production schedules to be taken into account.

REFERENCES

[1] International Journal of Forecasting, Official Publication of the International Institute of Forecasters, Interim Editor-in-Chief: Esther Ruiz, ISSN: 0169-2070, Elsevier B.V., www.journals.elsevier.com/international-journal-of-forecasting

[2] Journal of Forecasting, Editor-in-Chief: Derek W. Bunn, ISSN: 1099-131X, John Wiley & Sons Ltd, https://onlinelibrary.wiley.com/journal/1099131x

[3] Makridakis, Spyros & Andersen, A & Carbone, R & Fildes, Robert & Hibon, Michele & Lewandowski, R & Newton, Howard & Parzen, Emanuel & Winkler, R. (1982). The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition. Journal of Forecasting. 1. 111 - 153. 10.1002/for.3980010202.

[4] Rob J Hyndman (2018) A brief history of forecasting competitions, Monash Business School, Department of Econometrics and Business Statistics Working Paper Series 03/19, https://www.monash.edu/business/ebs/research/publications/ebs/wp03-2019.pdf

[5] Makridakis, Spyros & Chatfield, Chris & Hibon, Michele & Lawrence, Michael & Mills, Terence & Ord, Keith & F. Simmons, LeRoy. (1993). The M2-competition: A real-time judgmentally based forecasting study. International Journal of Forecasting. 9. 5-22. 10.1016/0169-2070(93)90044-N.

[6] Makridakis, Spyros; Hibon, Michele. The M-3 Competition: results, conclusions, and implications. International Journal of Forecasting. 16 (4): 451–476. doi:10.1016/S0169-2070(00)00057-1.

[7] Assimakopoulos, Vassilis & Nikolopoulos, Konstantinos. (2000). The theta model: A decomposition approach to forecasting. International Journal of Forecasting. 16. 521-530. 10.1016/S0169-2070(00)00066-2.

[8] Hyndman, Rob & Billah, Baki. (2003). Unmasking the Theta method. International Journal of Forecasting. 19. 287-290. 10.1016/S0169-2070(01)00143-1.

[9] L. Goodrich, Robert. (2000). The Forecast Pro methodology. International Journal of Forecasting. 16. 533-535. 10.1016/S0169-2070(00)00086-8.

[10] Makridakis, Spyros & Spiliotis, Evangelos & Assimakopoulos, Vassilis. (2018). The M4 Competition: Results, findings, conclusion and way forward. International Journal of Forecasting. 10.1016/j.ijforecast.2018.06.001.

[11] Smyl, Slawek. (2018). Exponential Smoothing - Recurrent Neural Network (ES-RNN) Hybrid Models. Uber Engineering. https://github.com/M4Competition/M4-methods/tree/slaweks_ES-RNN

[12] S. Ritter, M. Seidler, “A decision support system for high quality forecasting based on Artificial Intelligence”, Proceedings of ICDSST – PROMETHEE DAYS 2018, Greece, 2018.

[13] S. Ritter, M. Seidler, M. Jüttner, additional data to this paper, https://github.com/ifesca/ICDSST-2019, 2019

Forecast Stock Market with ARIMA, SVR, and GARCH

Xin James He
Fairfield University
1073 North Benson Road, Fairfield, CT 06824, USA
xhe@fairfield.edu

ABSTRACT

This research tries to forecast the U.S. stock market by comparing two major forecasting models: the autoregressive integrated moving average (ARIMA) and support vector regression (SVR). In addition, we deploy the generalized autoregressive conditional heteroscedasticity (GARCH), which is often used for financial times series that exhibit time-varying volatility or non-constant variance, to predict the volatility of the U.S. stock market. The main purpose of this research is to provide both analytical and computational insights for practitioners to select the right model for the right occasion in the marketplace. We find that while the ARIMA and SVR models can often lead to more accurate stock price forecasts, GARCH models can be used for financial market volatility predictions. Specifically, the SVR, as a machine learning model, performs better than the ARIMA models for stock price estimates, but it has a potential of overfitting and does not provide a list of parameters as in the case of ARIMA. The GARCH model becomes attractive to predict percentage returns and volatility when the financial market is experiencing significant ups and downs and sometimes frantic periods of trading. Managerial implications and computational insights are also discussed.

Keywords: Stock Forecast, Time Series, ARIMA, Machine Learning, SVR, GARCH

INTRODUCTION

Forecasting the stock market correctly has been an interesting subject both in academia and for practitioners, but few seemed to predict the sudden steep drops of the U.S. stock market in the month of December of 2018, which was considered as the second worst December in the history of the stock market performance. Ironically, the situation reversed so quickly in the following month, January 2019, when the US stock market rose to be the best January stock market in the last 30 years. While some experts believe it is impossible to predict the financial market under the Efficient Market Hypothesis, which implies a random walk model for stock prices, investment bankers and others try to use various computer models to forecast the stock prices and crude oil prices. In this research, we forecast the daily Dow Jones Industrial Average (DJI), S&P 500, and NASDAQ prices from Yahoo Finance over the time period 1986 - 2018 by comparing two commonly used forecasting models: the autoregressive integrated moving average (ARIMA) and support vector regression (SVR). In addition, we use the generalized autoregressive conditional heteroscedasticity (GARCH), which is often used for financial times series that exhibit time-varying volatility or non-constant variance, to predict the percentage returns and volatility of these indices. The main purpose of this research is to provide both analytical and computational insights for practitioners to select the right forecasting model for the right occasion in the marketplace by measuring the root mean squared error (RMSE) and mean absolute error (MAE). We find that while the ARIMA and SVR models can often lead to more accurate stock price forecasts, the GARCH model becomes attractive to predict percentage returns and volatility when the financial market is experiencing significant ups and downs and sometimes frantic periods of trading. Specifically, the SVR, as a machine learning model, performs better than the ARIMA for stock price estimates. However, since the SVR model, unlike ARIMA or other statistical models, does not provide a list of parameters to forecast for the future stock prices, it is often considered as a black box and has a potential of overfitting. Managerial implications and computational insights are also discussed.

This paper is organized as follows. Section 2 reviews the literature in dealing with financial market time series forecasting methods. Section 3 discusses research methodology in terms of data collections and analytical tools. Section 4 compares advanced time series ARIMA models with machine learning SVR and volatility prediction GARCH models. Finally, section 5 offers concluding remarks of this research.

LITERATURE REVIEW

Forecasting models for stock market prices can be divided into three major categories: advanced time series ARIMA, artificial intelligence or machine learning models SVR, and GARCH models [1]. Traditional time series models such as simple moving average (MA) and simple exponential smoothing (SES) are the most commonly used forecasting methods for time series data for U.S. government statistics and stock Wall Street prices. Advanced ARIMA models are the most prominent time series methods, in which autocorrelation function (ACF) and partial autocorrelation function (PACF) are used to help select data driven model parameters [2]. When it is done correctly, ARIMA models can provide very accurate forecasting results, especially for short-term time series data. MA, SES, and ARIMA models are often used as benchmarks to measure forecasting accuracy on financial markets against more complex machine learning models due to their easy to analyze and widespread applications.

SVR and artificial neural network (ANN) have been widely used to forecast crude oil and stock prices [3]. In this research, we exclude ANN because it requires a set of input variables in order to come up with an output variable through a complex function [4]. SVR is a special case of support vector machines (SVM), where SVM is a type of learning machine implementing the structural risk minimization inductive principle on a limited number of learning patterns. SVR computes a linear regression function in a high dimensional space where the input data are mapped via a nonlinear function. However, a major drawback of the SVR analysis is that it is difficult to interpret the process in meaningful statistical or business perspectives due to the fact that it relies on a high dimensional space via a nonlinear function.

Xie, Yu, Xu, and Wang (2006) find that SVR outperforms ARIMA based on monthly spot prices of West Texas Intermediate (WTI) crude oil from January 1970 to December 2004 [5]. However, the existing literature is very far from any consensus about a reliable forecasting model regarding crude oil and stock prices.

Engle (1982) introduced the ARCH and the GARCH models, which specifically allow for changing conditional variance [6]. The GARCH model requires joint estimation of the conditional mean model and the variance process. The GARCH (p, q) model is given in Eq.(1) below

$$R_t = \mu + \varepsilon_t \qquad \text{(Mean Equation)}$$
$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \qquad \text{(Variance Equation)} \tag{1}$$

where p is the order of GARCH while q is the order of ARCH process. Error, ε_t, is assumed to be normally distributed with zero mean and conditional variance, σ_t^2. R_t is the return, with the expect mean value μ to be positive and small. In addition, the parameters $\omega > 0$, $\alpha > 0$, and $\beta > 0$; $\alpha + \beta < 1$.

In this research, we compare the forecasting results of ARIMA and SVR on daily U.S. stock indices DJI, S&P 500, and NASDAQ during the period 1986 through 2018 to determine which model performs better in terms of RMSE and MAE. Moreover, we deploy the GARCH model to analyze the daily volatility of these indices during the same time period.

RESEARCH METHODOLOGY

We collect the daily U.S. stock market indices from Yahoo finance on DJI, S&P 500, and NASDAQ for the period January 1986 through December 2018. In this research, we focus our attention on model development and comparison so that we use all the data as he training data set, with no holdout data.

Figure 1 illustrates the time series of the entire DJI Index data January 1986 through December 2018. It is seen from Figure 1 that DJI Index prices fluctuate significantly, with a general trend going upward and also with periodic corrections and down turns. To meet the stationarity assumptions for ARIMA, we take the first order differencing the DJI Index daily prices as shown in Figure 2. Figure 3 shows the DJI daily percentage changes, which is to be used for the GARCH model.

It is interesting to note in Figures 2 that the price fluctuation increases toward the end of data set (i.e., around December 2018), whereas the percentage changes in Figure 3 were much more significant at the beginning of the data set (i.e., around October 1987). This is true for S&P 500 in Figures 4 - 6, and for NADDAQ in Figures 7 - 9. Consequently, the first order differencing as shown in Figs. 2, 5, and 8 can be considered as stationary for ARIMA models, with a constant mean and a roughly constant variance, whereas the percentage daily returns, $100(Y_t - Y_{t-1})/Y_{t-1}$, in Figures 3, 6, and 9 are commonly used for GARCH models.

Figure 1. DJI Index Daily Prices

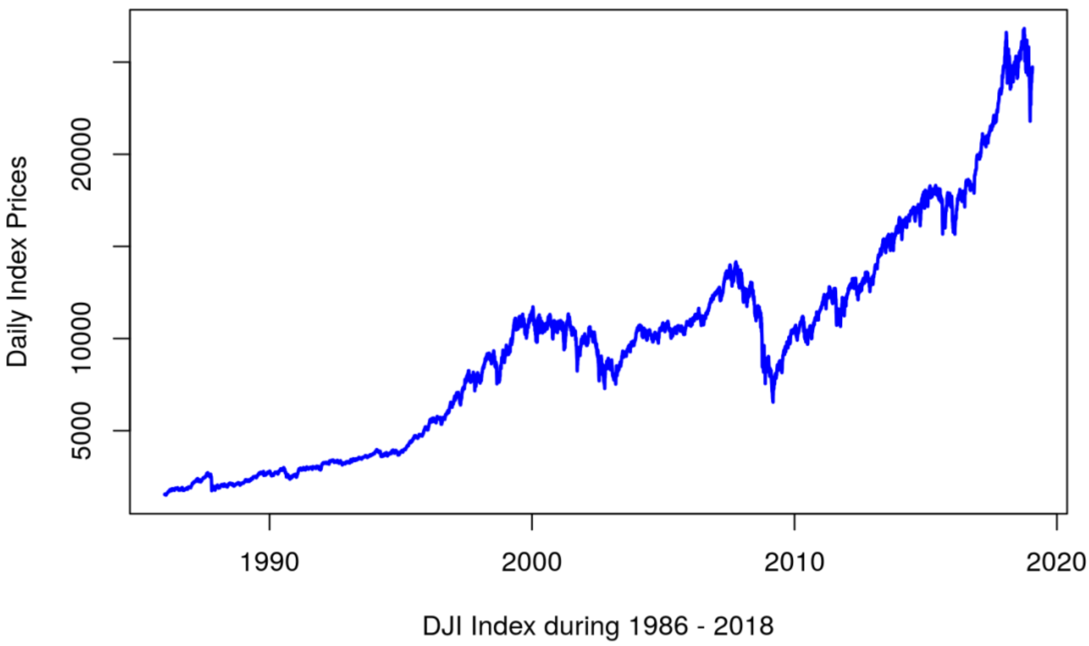


Figure 2. 1st Order Differencing on DJI

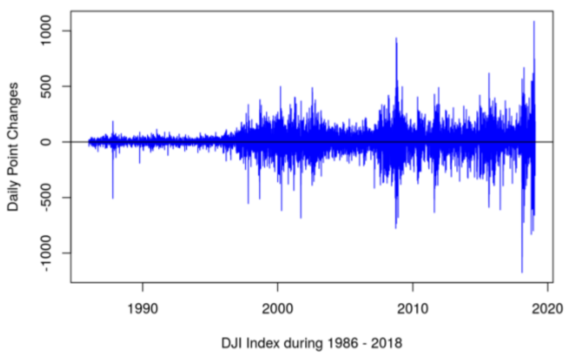
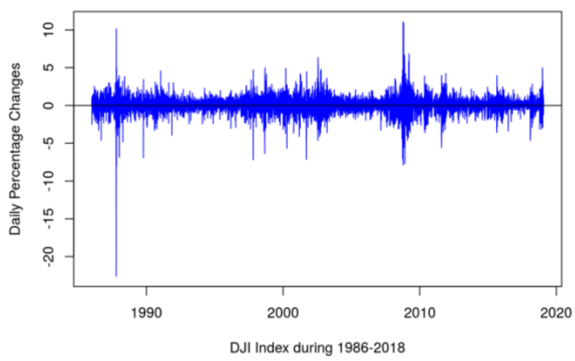


Figure 3. DJI Index Daily Return Data



We first use R to fit the ARIMA (p, d, q) models to determine the best parameters of p, d, and q by minimize the MAE and RMSE, and by the checking the stationarity and corresponding ACF & PCAF. We find that ARIMA (1, 1, 1) models without constant perform among all the models we tested for all three indices, DJI, S&P 500, and NASDAQ.

Table 1 shows the ARIMA (1, 1, 1) results of the DJI, S&P 500, NASDAQ prices, where Eq.(2) can be used to forecast future index prices.

Eq.(2) can be used to forecast future index prices.

$$\hat{Y}_t = (1+\phi)Y_{t-1} - \theta\epsilon_{t-1} \qquad \text{where } \epsilon_{t-1} = Y_{t-1} - \hat{Y}_{t-1} \qquad (2)$$

Table 1. ARIMA (1, 1, 1) Results 1986 - 2018

ARIMA (1,1,1)	DJI Daily Prices	S&P 500 Daily Prices	NASDAQ Daily Prices
AIC	102099.9	66283.85	84221.96
ϕ	0.6161	0.7153	0.7485
Std. Error (ϕ)	0.1887	0.0814	0.1003
θ	-0.6574	-0.7618	-0.7699
Std. Error (θ)	0.1799	0.0758	0.0970
RMSE	110.6251	12.90163	37.84711
MAE	68.06713	8.01276	21.50441

We then use R to fit the SVR models by selecting the appropriate parameters of cost, gamma, and epsilon. To avoid potential overfitting, we use the default epsilon = 0.1 and unscaled original daily indices for the period 1986 through 2018. Having tested numerous combinations of cost and gamma values, we narrow our search range to gamma = 0.1 – 0.3 and with a range of costs for different index. Then we use auto-tune in R package to come up with the optimal combination for costs ranging from 4000 to 6000 for DJI Index for example. Table 2 represents an auto-tune heat map produced by R for DJI Index prices, which indicates that the best performance for the SVR model lies in the upper left corner of the heat map.

Table 2. SVR Auto-Tune Heat Map on Cost and Gamma on DJI Index

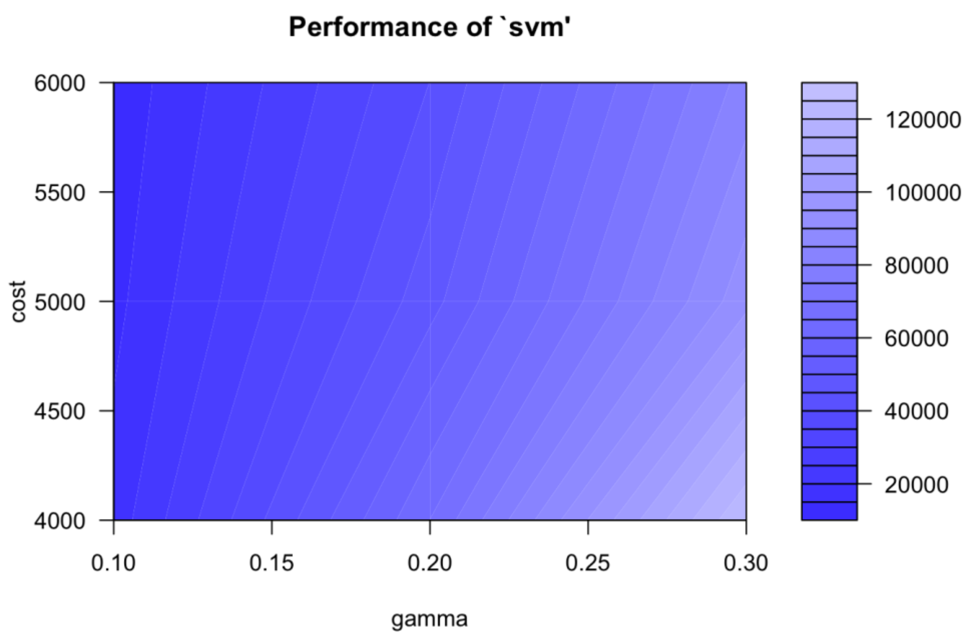


Table 3 compares ARIMA (1, 1, 1) models against SVR models for all three index prices in terms of RMSE and MAE. It is seen from Table 3 that SVR models provide more accurate results, but unlike the ARIMA (1, 1, 1) model in Table 1 with a list of parameters, an SVR model does not produce such parameters as in Eq.(2). Consequently, it cannot be used to forecast for future index prices.

Table 3. Index Prices Models 1986 – 2018

	Models	RMSE	MAE
DJI Daily Prices	**ARIMA (1, 1, 1)**	110.6251	68.0671
	SVR (c=6000, γ=0.1**)**	87.5687	23.3366
S&P 500 Daily Prices	**ARIMA (1, 1, 1)**	12.90163	8.01276
	SVR (c=600, γ=0.1**)**	11.18663	3.176736
NASDAQ Daily Prices	**ARIMA (1, 1, 1)**	37.84711	21.50441
	SVR (c=2000, γ=0.1**)**	29.94720	7.569939

CONCLUDING REMARKS

In this research, we forecast daily stock index prices of DJI, S&P 500, and NASDAQ using ARIMA, SVR, and GARCH models for both price prediction and volatility estimation with the following concluding remarks. First, the advanced autoregressive integrated moving average such as ARIMA (1, 1, 1) can offer more accurate forecasting results for most time series data, with reasonable computational complexity, and its model parameter(s) as shown in Eq.(2) can be used to forecast for the future or for the testing data. Second, while it can offer about the same forecasting accuracy as that of the ARIMA (1, 1, 1) model, the machine learning support vector regression such as SVR is not only computationally the more complex among all the forecasting models studied in this research, but also has the potential of model overfitting due to the fact that there are too many parameters to train the model: cost, gamma, and epsilon. Third, an SVR model cannot be used to test the model accuracy on the testing data the same way as in an ARIMA model since it does not provide a list of model parameters as in Eq.(2) of the ARIMA model, which also makes the economic or business interpretation very difficult.

REFERENCES

[1] Behmiri and Manso (Jan. 2013). Crude Oil Forecasting Techniques: A Comprehensive Review of Literature, SSRN Electronic Journal.

[2] Ord, Keith, Robert Fildes, and Nikolaos Kourentzes (2017). Principles of Business Forecasting, Wessex Learning, USA.

[3] Sehgal, Neha and Pandey, Krishan K. (2015). Artificial Intelligence Methods for Oil Price Forecasting: A Review and Evaluation. Energy Systems, 6, 479-506.

[4] Haidar, I., Kulkarni, S., Pan, H. (2008). Forecasting model for crude oil prices based on artificial neural networks. In: International Conference on Intelligent Sensors, Sensor Networks and Information Processing, (ISSNIP 2008), 103–108. IEEE.

[5] Xie, W., Yu, L., Xu, S., Wang, S. (2006). A new method for crude oil price forecasting based on support vector machines. International Conference on Computational Science

(Part IV), 444-451

[6] Engle, R F (1982). "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation", Econometrica, 50 (4), pp. 987-1008.

How early can we predict student's MOOC performance? – A DSS for course stopout mitigation

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

ABSTRACT

Due to high attrition rates in MOOCs problem of identification who will fail an exam, which is called stopout, has become of immense importance for course designers and course makers. MOOC systems generate a lot of data, however, because of the huge amount of data, classical statistical methods are often not applicable. Therefore, stopout can be detected using machine learning algorithms. Based on predictive model predictions a course designer can make a decision support system which will act as an early warning system and make intelligent decisions motivating the student to invest extra efforts in order to pass an exam. In this paper lasso and ridge logistic regression are created for 22 courses offered by Open University, containing over 20,000 students. As descriptors of the student, interaction with a virtual learning environment are used. Besides answering the question if we can predict who will fail an exam, we also give an answer to the question of how early we can make that prediction. Results show that before the start of the course predictive performance is fairly good with AUC 0.658, and after six months of course AUC rises to 0.931.

Keywords: Educational Data Mining, Logistic Regression, Stopout prediction

INTRODUCTION

The higher education sector is experiencing a massive change with learning materials shifting from printed media and in-house learning environments to virtual learning environments with interactive materials available on the Internet [[9]]. One kind of virtual learning environment called Massive Open Online Courses (MOOCs) became popular in recent years. One can access learning materials from world-wide renown Universities and lecturers, often provided free of charge. Additionally, one can choose from a wide variety of subjects, at ordering and pace one chooses [[2]]. Because of that a lot of students are attracted to this model of education.

However, this variety of learning subjects and ease of access to learning materials provided new problems which are not seen in traditional learning environments. Namely, the majority of the students do not intent to complete a course. Majority of them are dropouts, meaning that they unenrolled from course materials before the formal end of the course. Although this is an important problem present in a modern learning environment it is hard to motivate the student to finish course examinations without sacrificing the quality of the course and/or course design. The other problem, probably more important, present a situation where a student was enrolled in the online course but hasn't earned enough points to receive a certificate. One can say that this is not necessarily a problem, because students may enroll for different reasons, i.e. access to learning materials. However, if the student interacted with the learning environment (i.e. listened to lectures, discussed in the forum and attempted assignments) and did not complete the course then further investigation is needed. For example, due to various reasons such as insufficient background knowledge, lack of time, course design or other students, one felt discouraged, frustrated, or bored which resulted in the student not finishing the course. In the literature review, one can find that this is a major problem MOOCs are facing.

A lot of effort is invested in analyzing how and why students are prone to dropout and failing to pass the exam, especially in MOOCs. This area is called Learning Analytics or Educational Data Mining [[10]]. There are multiple reasons why many classical statistical approaches fail, according to [[12]]. First, there is a lack of "negative samples", i.e. only several percents of students obtain a certificate of completion (the majority of the students are considered as a dropout). Also, this may result in biasing of the machine learning models toward the majority of the class. Another problem is a huge amount of unstructured data. The third problem is data variance, which is the consequence of MOOC's self-paced learning pedagogy (i.e. student learn at different pace, with different learning habits). Because of the above-mentioned machine learning approach with a lot of preprocessing is needed in order to obtain the usable predictive model. [[12]]

In this paper, we will utilize Open University Learning Analytics dataset [[8]] with the goal of predicting whether a student will fail to pass the exam on 22 courses originating from Open University. Namely, the goal of the paper is to predict how early we can predict whether a student will fail to pass the exam. The predictive model will act as an early warning system which will allow the decision maker to control the level of activation. For that purposes, we use data originating from the virtual learning environment, i.e. logs of the interaction of students with learning materials. An additional question we answer is how early we can predict whether a student will fail an exam.

The following of the paper is organized as follows. In Section 2 related work is presented. In Section 3 we present methodology. We will present data used in this research, followed by the experimental setup and evaluation of the predictive model. In Section 4 we provide results and interpretation of results, while in Section 5 we conclude the paper.

RELATED WORK

From a historical point of view, first MOOC was created by George Siemens and Stephen Downs and was called "Connectivism and Connective Knowledge" in 2008 and it attracted 25 students who participated paying fees for the course, and over 2,000 students enrolling to the course free of charge. Nowadays, through virtual learning environments such as Coursera, EdX or Udacity one can find courses with over 1,000,000 enrolled learners coming from over 190 countries [[13]].

The term dropout or stopout is considered as the last event student participated in, such as quiz assignment [[15]] or watching a video [[1]]. However, it is also commonly adopted that stopout is identified as a situation when a student does not earn a certificate of accomplishment within a course [[7], [6]]. The last definition of the stopout is the one we adopted, however instead of calling it stopout we will call it failing to pass the exam.

Most often predicting whether a student will fail an exam is set as a binary classification problem. For example, Jiang et al [[7]] used performance on assignment from the first week. It has been shown that performance on the first-week assignment is enough to discriminate between students with a certificate of accomplishment and certificate of accomplishment with distinction with AUC of 0.947 and also between students with a certificate of accomplishment and students which failed to pass the exam with AUC 0.851. Similarly, activities on quizzes and assignments, as well as activity on the forum can be used to predict performance on final exam [[11]]. Due to the fact that identification of poorly performing students should be done as early as possible when no student behavior data is available, one can find predictive models which utilize only demographic data such as gender, occupational status, age and self-ascribed motivation [[14]]. However, most often clickstream data, from which features about student interaction from virtual learning environment are extracted are most commonly used. Those features are most often regarding how often student interact with video learning materials, activity on a discussion forum, how much time the student spent on the page etc. [[15]]. Also, as an important features one can find grades achieved on assignments and exams [[7], [3], [4]].

It is noted that as a measure of performance most often Area Under the Receiver Operating Characteristics Curve (AUC) measure is used.

METHODOLOGY

Data and Algorithm

In this paper, we utilized Open University Learning Analytics Dataset [[8]] which contain data originating from Open University student from 2013 and 2014. This includes besides demographic data also interaction data of students with University's virtual learning environment.

Logistic regression is a statistical learning algorithm commonly used for the stopout problem [[15], [7], [6]]. It is used for classification in which there are one or more independent features that determine an outcome. We can define logistic regression as a classifier that models the probability of dependent binary features y given a set of independent features X using logit model [[5]].

In order to prevent overfitting, one can find extensions of $L(\theta)$. Most popular and most commonly used are Lasso and Ridge logistic regression. Lasso adds L1 regularization term in $L(\theta)$ which forces coefficients of logistic regression to zero, i.e. lowers the number of features needed to explain the problem at hand. This way complexity of the problem is reduced. Ridge logistic regression adds L2 regularization which forces coefficients of logistic regression to be lower in general. Both regularization terms are used to prevent model to explain random noise or error. However, regularization terms introduce hyper-parameter λ which needs to be optimized. [[16]]

Experimental Setup

In order to answer both research questions, we trained and tested predictive models in seven different experimental settings. We create predictive models based on demographic and personal features such as gender, disability, number of previously finished courses, age, a region of residence of the student, index of multiple deprivations, and level of education. Those are attributes available for prediction model before the course started. The second experimental setup consists of demographic and personal features and activities of the student on assessment and virtual learning environment in the first month of course. Those features are, among others, number of assessment student tried to pass, the average score on assessments, number of passed student assessment, number of activities student interacted with etc. The third experimental setup consists of features of previous experimental setup and activities of the student on assessment and virtual learning environment in the second month of course. The experimental setups are created in the same manner up to six months which consists of demographic and personal features and activities of the student on assessment and virtual learning environment in the first, second, third, fourth, fifth and sixth month of the course.

Models are evaluated using AUC because it is commonly used in educational data mining applications and specifically for the problem at hand.

In order to obtain valid results and comparisons, the model is validated using ten-fold cross-validation, while regularization hyper-parameter λ is obtained using inner ten-fold cross-validation.

RESULTS

In Table 1 we present obtained results for different experimental setups. Rows present different experimental setups, while columns present learning algorithm used. Namely, those are Lasso Logistic Regression (Lasso LR) and Ridge Logistic Regression (Ridge LR). Values inside Table 1 present average AUC with standard deviation obtained using ten-fold cross-validation. We note that later experimental setups include features from previous experimental setups.

Table 6: Performances of Logistic Regression for different experimental setups

Experimental Setup	Lasso LR	Ridge LR
Demographic and Personal Features	0.658 ± 0.014	0.657 ± 0.014
Activities in first month	0.750 ± 0.009	0.748 ± 0.009
Activities in second month	0.805 ± 0.003	0.800 ± 0.004
Activities in third month	0.832 ± 0.003	0.826 ± 0.004
Activities in fourth month	0.872 ± 0.005	0.866 ± 0.005
Activities in fifth month	0.906 ± 0.004	0.900 ± 0.005
Activities in sixth month	0.931 ± 0.004	0.925 ± 0.005

As we can observe from Table 1 performances of predictive model increases as more information is added. Namely, by adding more features which present activities of the student on assessment and interaction with virtual learning environment predictive performance are better. It is interesting to note that demographic and personal features can discriminate performance on the exam. One can also notice that Lasso logistic regression is consistently better compared to Ridge logistic regression, but the difference is on the third decimal place.

AUC curves are presented in Figure 2. As already stated, adding more information provide curves with higher areas. The best performing predictive model (activities in the sixth month) is presented in red color. For the best performing model, we can state that over 60% of students which will fail an exam can easily be identified without single false positive for both Lasso and Ridge logistic regression. The percentage is lowering for earlier every month and for the predictive model using information from the first three months around 37% of students which will fail an exam can be identified without a single false alarm. Similarly, if decision maker sets a threshold that 80% of students who will fail an exam should be identified then the best performing model will have around 16% of false alarms, predictive model using five months data around 21% of false alarms, while demographic and personal feature-based model would have around 78% of false alarms.

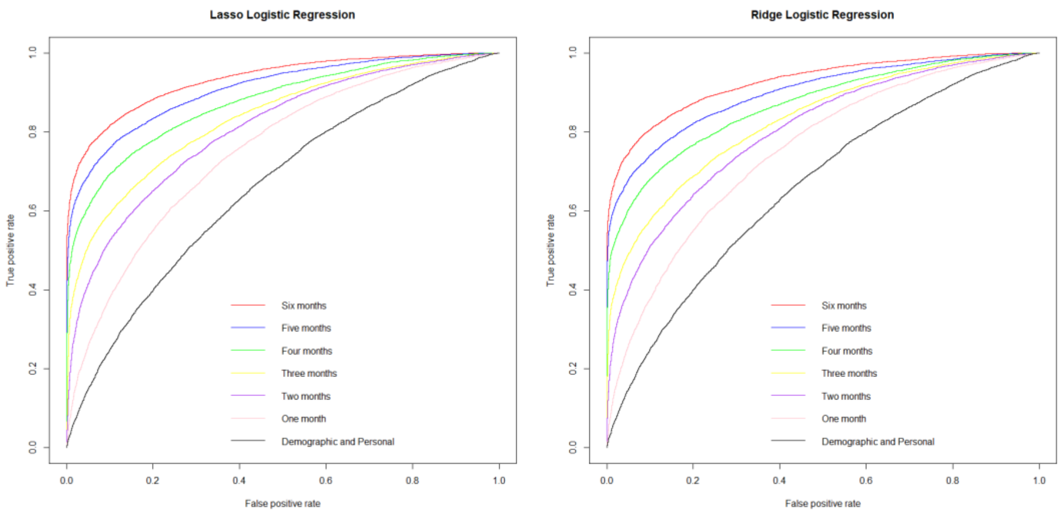


Figure 13: AUCs of Lasso and Ridge Logistic Regression

CONCLUSIONS

Education data mining and application of machine learning techniques in educational settings presents an interesting research area which requires not only technical skills of data analyst but also pedagogical and communicational skills in order to make effective decisions. In this paper, we propose lasso and ridge logistic regression models for prediction of student's stopout. In order to answer the question of how early we can make a prediction we developed seven experimental setups. First experimental setup does not require any behavioral data of the student. Namely, we used only demographic and personal data. Results were interesting with AUC 0.658 and require an additional inspection. Further experimental setups included student interaction with a virtual learning environment, such as results of assessments and interaction with learning materials. Namely, we extracted features which explain interaction with the learning environment in the first month, second month, third month, fourth month, fifth month and sixth month of the course. Results suggest that adding more data, i.e. using data from the latter part of the course improve the predictive performance of the model. More specifically, AUC increases from 0.750 after the first month of the course to AUC 0.931 after six months of the course.

As a part of future research, we plan to explore obtained coefficients from logistic regression in order to answer the question of why students fail an exam. By understanding the reasons why more personalized intervention can be created. Also, we would like to improve the predictive performance of the proposed model by utilizing multi-task logistic regression models. More specifically, we would like to create a predictive model for each course but with knowledge sharing between courses.

REFERENCES

[1] Allione, G., & Stein, R. M. (2016). Mass attrition: An analysis of drop out from principles of microeconomics MOOC. The Journal of Economic Education, 47(2), 174-186.

[2] Chengjie, Y. U. (2015). Challenges and changes of MOOC to traditional classroom teaching mode. Canadian Social Science, 11(1), 135-139.

[3] Dalipi, F., Imran, A. S., & Kastrati, Z. (2018, April). MOOC dropout prediction using machine learning techniques: Review and research challenges. In Global Engineering Education Conference (EDUCON), 2018 IEEE (pp. 1007-1014). IEEE.

[4] Gardner, J., Brooks, C., Andres, J. M., & Baker, R. S. (2018, December). MORF: A Framework for Predictive Modeling and Replication At Scale With Privacy-Restricted MOOC Data. In 2018 IEEE International Conference on Big Data (Big Data) (pp. 3235-3244). IEEE.

[5] Hastie, T., Tibshirani, R., & Wainwright, M. (2015). Statistical learning with sparsity: the Lasso and generalizations. CRC press.

[6] He, J., Bailey, J., Rubinstein, B. I., & Zhang, R. (2015, January). Identifying At-Risk Students in Massive Open Online Courses. In AAAI (pp. 1749-1755).

[7] Jiang, S., Williams, A., Schenke, K., Warschauer, M., & O'Dowd, D. (2014, July). Predicting MOOC Performance with Week 1 Behavior. In Educational Data Mining, 2014.

[8] Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open University Learning Analytics Dataset. Scientific Data, 4, 170171.

[9] Muilenburg, L. Y., & Berge, Z. L. (Eds.). (2016). Digital badges in education: Trends, issues, and cases. Routledge.

[10] Papamitsiou, Z., & Economides, A. A. (2014). Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence. Journal of Educational Technology & Society, 17(4).

[11] Ramesh, A., Goldwasser, D., Huang, B., Daume III, H., & Getoor, L. (2014, June). Learning latent engagement patterns of students in online courses. In Twenty-Eighth AAAI Conference on Artificial Intelligence.

[12] Romero, C., & Ventura, S. (2017). Educational Data Science in Massive Open Online Courses. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 7(1), e1187.

[13] Sanchez-Gordon, S., & Luján-Mora, S. (2016). How could MOOCs become accessible? The case of edX and the future of inclusive online learning.

[14] Stein, R., & Allione, G. (2014). Mass attrition: An analysis of drop out from a Principles of Microeconomics MOOC.

[15] Taylor, C., Veeramachaneni, K., & O'Reilly, U. M. (2014). Likely to stop? Predicting stopout in Massive Open Online Courses. arXiv preprint arXiv:1408.3382.

[16] Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: a retrospective. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73(3), 273-282.

Developing a Flexible Decision Support System for Air Cargo Forwarding Problems under Uncertainty

Lin Zhu and Yue Wu
Southampton Business School, University Road, Southampton, UK, SO17 1BJ
y.wu@soton.ac.uk

Abstract. This research is concerned with air cargo forwarding problems under uncertainty. Cargoes need to be transported from different regions to different destinations via a hub. Air forwarders need to make a decision about the number of containers to be booked for the regions and hub in advance before accurate customers' information becomes available. A decision support system that is based on a robust model is developed to help forwarders to book air containers in advance under uncertain information. The decision support system can handle the situation where cargo delay is allowed as long as it is profitable to do so. Computational results show the effectiveness and flexibility of the decision support system.

Keywords: Air Container, Air Cargo, Robust Model.

Introduction

Transportation by air is the quickest way to deliver items over long distances. Transit time has been reduced from nearly 60 days for shipping, to just a day or two, because of the massive sizes of cargo planes, the high frequency of numerous airlines, and the presence of many airports in major cities around the world. Air transport is, however, very expensive comparing with other transport modes. Due to the rapid growth in airline business, airline hubs have been considered in recent years to facilitate more efficient use of scarce air transportation resources to reduce costs. When air cargo is forwarded via hubs, goods are brought from the source region by air flights to the hub where they are consolidated and packed to other aircrafts, which then convey the goods to their ultimate destinations. Therefore, the questions of how to book air containers, in the region and hub in order to transport the products to distant markets on time have become a challenge to decision makers.

This research aims to develop a decision support system to help air forwarders to book air containers in order that cargoes from different regions could be shipped to a hub, where they are repacked and consolidated before leaving for various destinations. The problem becomes more complicated when customer information is uncertain in decision-making process. Every type of cargo has its own weight and volume. Cargo cannot be divided which means that each cargo must be loaded into one container. Air forwarders need to book air containers in all regions and hub. Notice that, in the hub, containers can come from three sources: one is from the regions, containers that can stay in use, called pre-used containers; another resource is from booking in advance, in which case they are called new containers; another resource is urgent booking on the shipping day. Using pre-used containers rather than new containers in the hub will get a discount from airlines. Containers are normally booked one week before the shipping date for a cheap rental price from airlines. The cargo quantity that customers provide is uncertain when booking. The forwarders do not want to wait until the actual shipment information is observed, which would normally happens on the date of shipping, as it is very expensive to require extra containers on the shipping day. Therefore, before the time of rental, accurate information for cargo information is uncertain. Forwarders have to make a decision about the quantities and types of containers in all regions and the hub, along with how to use the pre-used containers in the hub. Then, after the realization of the uncertainty, if the containers that have been booked cannot hold all cargoes, additional containers are required with a high cost. On the other hand, if too many containers have been ordered, redundant containers have to be returned to the airlines with a cost due to breaking of a contract with airlines. Therefore, the rental cost consists of two parts: the cost of using the containers and the cost of requiring extra containers and returning unused containers on the day of shipping. The cost of using a container is based on a fixed charge, plus a variable charge that depends on the total cargo weight that the container holds. In addition to the rental cost, there is a cost, called penalty cost that occurs when cargoes are not shipped on the day of shipping. The forwards could decide not to ship some cargoes if it is too expensive. However, it will cause the penalty cost because of breaking a contract to customers.

Literature Review

The present literature contributions on the problems of container loading mainly focus on sea container loading. Over the last few years, a lot of publications have been released discussing the problems faced when loading air cargo. Cargo container loading plans is conducted in [1], using the operations of the carrier FedEx. Airlines looking for a best possible baggage limit policy, whilst the goods were carried in the remaining aircraft belly space along with customers' luggage, has been identified as a new difficulty in [2]. In [3], there are two models: a single flight and a sequential multi-flight. Both models rely on the uncertain capacity restriction from the point of view of the freight forwarders who are tasked with determining the booking amounts for spot and contract markets. The models also consider the problems of capacity booking. A mixed integer linear programming model is formulated in [4] to assist logistic managers in making the necessary decisions when they are dealing with containerization of problems encountered while loading air cargo: precisely, when leasing air containers from air carrier service providers, as well as how to optimally load air cargo into their required containers. However, the model does not address the uncertainty challenge when the accurate information cannot be obtained at the time of booking. In order to solve this problem, a stochastic mixed 0-1 model is proposed in [5] for a dual-response forwarding system for booking air containers and determining how cargoes are loaded in the containers simultaneously under uncertainty. [6] goes further and extends the stochastic model to a robust model for addressing similar challenges in which cargo is permitted to be shipped at a later date. The robust model uses a quantitative approach to measuring the trade-off between the costs involved and the risks expected. Only few researchers focus on air container booking problems, especially considering to the addition of a hub to consolidate air freights. There is also no paper contributing on the air container booking problems for multi-flight via a hub under uncertainty.

Developing a Robust Model for Decision Support System

Container Rental Costs

The cost of using a container includes two parts, fixed cost and variable cost. Once a container is selected, the fixed cost needs to be paid. If the weight of cargo loaded into the container exceeds some given values, a variable cost will be incurred. These values are called break point. Let $\{1, 2, \cdots, K\}$ be K break points. The K break points divide the container weight limit into K intervals. The unit variable cost in a k interval is charged at a slope rate δ_k. In this study, the air carriers provide six break points, a_1, a_2, a_3, a_4, a_5 and a_6, for each type of container, $j = 1, 2, \cdots, n$. n represents the maximum number of air cargo types. w_j denotes the weight of cargo j and y_j is the quantity of air cargo j loaded into this container. The variable cost is the same as in [5], which is a piecewise function defined as:

$$c = \begin{cases} 0 & \sum_{j=1}^{n} w_j y_j \in (0, a_1] \\ \delta_2(\sum_{j=1}^{n} w_j y_j - a_1) & \sum_{j=1}^{n} w_j y_j \in (a_1, a_2] \\ \delta_2(a_2 - a_1) & \sum_{j=1}^{n} w_j y_j \in (a_2, a_3] \\ \delta_2(a_2 - a_1) + \delta_4(\sum_{j=1}^{n} w_j y_j - a_3) & \sum_{j=1}^{n} w_j y_j \in (a_3, a_4] \\ \delta_2(a_2 - a_1) + \delta_4(a_4 - a_3) & \sum_{j=1}^{n} w_j y_j \in (a_4, a_5] \\ \delta_2(a_2 - a_1) + \delta_4(a_4 - a_3) + \delta_6(\sum_{j=1}^{n} w_j y_j - a_5) & \sum_{j=1}^{n} w_j y_j \in (a_5, a_6] \end{cases} \quad (1)$$

We can simplify (1) by adding two new variables g_k and z_k. g_k represents the cargo weight in the range $(a_{k-1}, a_k]$ in the container. $z_k = \begin{cases} 1 & \text{if } g_k > 0 \\ 0 & \text{otherwise} \end{cases}$. Then the variable cost can change to $\sum_{k=1}^{K} \delta_k g_k$ by adding some constraints:

$$\sum_{k=1}^{K} g_k = \sum_{j=1}^{n} w_j y_j \tag{2}$$

$$g_k \leq z_k (a_k - a_{k-1}), \quad k = 1, 2, \cdots, K \tag{3}$$

$$g_k \geq z_{k+1} (a_k - a_{k-1}), \quad k = 1, 2, \cdots, K \tag{4}$$

Equations (3) and (4) ensure that g_k cannot be positive unless the range $(a_{k-1}, a_k]$ is fully occupied by the cargo weight.

Notations

Indices

i types of containers $(i = 1, 2, \cdots, m)$;
j types of cargoes $(j = 1, 2, \cdots, n)$;
r regions $(r = 1, 2, \cdots, R)$;
d destinations $(d = 1, 2, \cdots, D)$;
s scenarios $(s = 1, 2, \cdots, S)$;
k numbers of breaking-points for type i container $(k = 1, 2, \cdots, K_i)$;
l numbers of type i container $(l = 1, 2, \cdots, L_i)$.

Parameters

v_j volume of a type j cargo;
w_j weight of a type j cargo;
V_i volume limit of type i container;
W_i weight limit of type i container;
a_{ik} weight of type i container in breaking-points k;
δ_{ik} the unit charge rate of type i container in the range $(a_{i(k-1)}, a_{ik}]$;
c_{ir}^0 fixed cost by renting a type i container in region r;
c_i^{h0} fixed cost by renting a type i container in hub;
q_{jsr} type j cargo quantity in scenario s in region r;
p_s probability of scenario s;
L_{ir} type i container available quantity in region r;
L_i type i container available quantity in all regions, which means $L_i = \sum_{r=1}^{R} L_{ir}$;
L_i^h type i container available quantity in hub;
c_{ir}^- / c_{ir}^+ the unit cost of requiring/returning type i containers on the day of shipping in region r;
c_i^{h-} / c_i^{h+} the unit cost of requiring/returning type i containers on the day of shipping in hub;
b_i the unit repacking cost of type i container in the hub;
θ the discount rate of fixed cost by using pre-used containers;
q_{jsd}^h quantity of type j cargo with destination d in scenario s in the hub.

Decision variables

o_{ir} number of type i container for booking in region r;
o_i^h number of type i container for booking in hub;
o_{isr}^- / o_{isr}^+ number of type i container required/returned in scenario s on the day of shipping in region r;
$o_{is}^{h-} / o_{is}^{h+}$ number of type i container required/returned in scenario s on the day of shipping in hub;
o_i^{hc} number of type i pre-used container for booking to continue to use in the hub;

$x_{ilsr} = \begin{cases} 1 & \text{if the } l\text{th container of type } i \text{ is selected in scenario } s \text{ in region } r \\ 0 & \text{otherwise} \end{cases}$;

y_{iljsrd} quantity of type j cargo with destination d loaded into the lth container of type i in scenario s in region r;

y^h_{iljsd} quantity of type j cargo with destination d loaded into the lth container of type i in scenario s in hub;

g_{ilksr} cargo weight distributed in the range $(a_{i(k-1)}, a_{ik}]$ inside the lth container of type i in scenario s in region r;

g^h_{ilksd} cargo weight distributed in the range $(a_{i(k-1)}, a_{ik}]$ inside the lth container of type i in scenario s with destination d in the hub;

$z_{ilksr} = \begin{cases} 1 & \text{if } g_{ilksr} > 0 \\ 0 & \text{otherwise} \end{cases}$;

$z^h_{ilksd} = \begin{cases} 1 & \text{if } g^h_{ilksd} > 0 \\ 0 & \text{otherwise} \end{cases}$;

$x^h_{ilsd} = \begin{cases} 1 & \text{if the } l\text{th type } i \text{ container with destination } d \text{ is selected in scenario } s \text{ in hub} \\ 0 & \text{otherwise} \end{cases}$

$x^{hc}_{ilsd} = \begin{cases} 1 & \text{if } l\text{th type } i \text{ pre}-\text{used container with destination } d \text{ is selected in scenario } s \text{ in hub} \\ 0 & \text{otherwise} \end{cases}$

y^{hc}_{iljsd} quantity of type j cargo with destination d loaded into the lth pre-used container of type i in scenario s in the hub;

g^{hc}_{ilksd} cargo weight distributed in the range $(a_{i(k-1)}, a_{ik}]$ inside the lth type i pre-used container with destination d in scenario s in the hub;

$z^{hc}_{ilksd} = \begin{cases} 1 & \text{if } g^{hc}_{ilksd} > 0 \\ 0 & \text{otherwise} \end{cases}$

A robust model

Model robustness for air cargo forwarding problems means that the air cargoes can be transported next week by adding a penalty cost ω_j per unit. The decisions in the first stage are o_{ir}, o^h_i and o^{hc}_i, where o_{ir} and o^h_i are the container booking decisions in regions and hub, and o^{hc}_i is the decision using pre-used containers in the hub. In the second stage, the decisions for urgent renting or returning containers on the shipping day are o^-_{isr}, o^+_{isr}, o^{h-}_{is} and o^{h+}_{is}, respectively. The decisions for loading air cargoes into the containers are y_{iljsrd}, y^h_{iljsd} and y^{hc}_{iljsd}. e_{jsrd} and e^h_{jsd} represent the unshipped cargos in regions and hob, respectively. The robust model is shown below:

$$\min \sum_{r=1}^{R} \left(M_r + \sum_{i=1}^{m} \sum_{s=1}^{S} p_s c^-_{ir} o^-_{isr} + \sum_{i=1}^{m} \sum_{s=1}^{S} p_s c^+_{ir} o^+_{isr} + \sum_{d=1}^{D} \sum_{j=1}^{n} \sum_{s=1}^{S} p_s \omega_j e_{jsrd}\right) + \sum_{r=1}^{R} \sum_{i=1}^{m} \sum_{l=1}^{L_{ir}} \sum_{s=1}^{S} b_{it} x_{ilsr} + \sum_{d=1}^{D} (N^c_d + N_d) + \sum_{i=1}^{m} \sum_{s=1}^{S} p_s c^{h-}_i o^{h-}_{is} + \sum_{i=1}^{m} \sum_{s=1}^{S} p_s c^{h+}_i o^{h+}_{is} + \sum_{d=1}^{D} \sum_{j=1}^{n} \sum_{s=1}^{S} p_s \omega^h_j e^h_{jsd} \quad (5)$$

subject to

$$M_r = \sum_{i=1}^{m} \sum_{l=1}^{L_{ir}} \sum_{s=1}^{S} p_s c^0_{ir} x_{ilsr} + \sum_{i=1}^{m} \sum_{l=1}^{L_{ir}} \sum_{k=1}^{K_i} \sum_{s=1}^{S} p_s \delta_{ik} g_{ilksr} \quad (6)$$

$$N^c_d = \sum_{i=1}^{m} \sum_{l=1}^{L_i} \sum_{s=1}^{S} p_s \theta c^{h0}_i x^{hc}_{ilsd} + \sum_{i=1}^{m} \sum_{l=1}^{L_i} \sum_{k=1}^{K_i} \sum_{s=1}^{S} p_s \delta_{ik} g^{hc}_{ilksd} \quad (7)$$

$$N_d = \sum_{i=1}^{m} \sum_{l=1}^{L^h_i} \sum_{s=1}^{S} p_s c^{h0}_i x^h_{ilsd} + \sum_{i=1}^{m} \sum_{l=1}^{L^h_i} \sum_{k=1}^{K_i} \sum_{s=1}^{S} p_s \delta_{ik} g^h_{ilksd} \quad (8)$$

$$o_{ir} = \sum_{l=1}^{L_{ir}} x_{ilsr} + o^+_{isr} - o^-_{isr} \quad (9)$$

$$o_i^h = \sum_{l=1}^{L_i^h} \sum_{d=1}^{D} x_{ilsd}^h + o_{is}^{h+} o_{is}^{h-} \quad (10)$$

$$\sum_{l=1}^{L_i} \sum_{d=1}^{D} x_{ilsd}^{hc} \leq \sum_{r=1}^{R} \sum_{l=1}^{L_{ir}} x_{ilsr} \quad (11)$$

$$\sum_{l=1}^{L_i^h} \sum_{d=1}^{D} x_{ilsd}^h \leq L_i^h \quad (12)$$

$$\sum_{l=1}^{L_i} \sum_{d=1}^{D} x_{ilsd}^{hc} = o_i^{hc} \quad (13)$$

$$\sum_{k=1}^{K_i} g_{ilksr} = \sum_{j=1}^{n} \sum_{d=1}^{D} w_j y_{iljsrd} \quad (14)$$

$$g_{ilksr} \leq z_{ilksr}\left(a_{ik} - a_{i(k-1)}\right) \quad (15)$$

$$g_{ilksr} \geq z_{il(k-1)sr}\left(a_{ik} - a_{i(k-1)}\right) \quad (16)$$

$$\textstyle\sum_{k=1}^{K_i} g_{ilksd}^h = \sum_{j=1}^{n} w_j y_{iljsd}^h \quad (17)$$

$$g_{ilksd}^h \leq z_{ilksd}^h\left(a_{ik} - a_{i(k-1)}\right) \quad (18)$$

$$g_{ilksd}^h \geq z_{il(k-1)sd}^h\left(a_{ik} - a_{i(k-1)}\right) \quad (19)$$

$$\sum_{k=1}^{K_i} g_{ilksd}^{hc} = \sum_{j=1}^{n} w_j y_{iljsd}^{hc} \quad (20)$$

$$g_{ilksd}^{hc} \leq z_{ilksd}^{hc}\left(a_{ik} - a_{i(k-1)}\right) \quad (21)$$

$$g_{ilksd}^{hc} \geq z_{il(k-1)sd}^{hc}\left(a_{ik} - a_{i(k-1)}\right) \quad (22)$$

$$\sum_{i=1}^{m} \sum_{d=1}^{D} \sum_{l=1}^{L_{ir}} y_{iljsrd} = q_{jsr} - e_{jsrd} \quad (23)$$

$$\sum_{i=1}^{m} \sum_{l=1}^{L_i^h} y_{iljsd}^h + \sum_{i=1}^{m} \sum_{l=1}^{L_i} y_{iljsd}^{hc} = q_{jsd}^h - \sum_{r=1}^{R} e_{jsrd} - e_{jsd}^h \quad (24)$$

$$\sum_{j=1}^{n} \sum_{d=1}^{D} v_j y_{iljsrd} \leq V_i x_{ilsr} \quad (25)$$

$$\sum_{j=1}^{n} \sum_{d=1}^{D} w_j y_{iljsrd} \leq W_i x_{ilsr} \quad (26)$$

$$\sum_{j=1}^{n} v_j y_{iljsd}^h \leq V_i x_{ilsd}^h \quad (27)$$

$$\sum_{j=1}^{n} w_j y_{iljsd}^h \leq W_i x_{ilsd}^h \quad (28)$$

$$\sum_{j=1}^{n} v_j y_{iljsd}^{hc} \leq V_i x_{ilsd}^{hc} \quad (29)$$

$$\sum_{j=1}^{n} w_j y_{iljsd}^{hc} \leq W_i x_{ilsd}^{hc} \quad (30)$$

$$o_{ir}, o_i^h, o_i^{hc}, o_{isr}^-, o_{isr}^+, o_{is}^{h-}, o_{is}^{h+}, y_{iljsrd}, y_{iljsd}^h, y_{iljsd}^{hc}, g_{ilksr}, g_{ilksd}^h, g_{ilksd}^{hc}, e_{jsrd}, e_{jsd}^h \text{ are nonnegative integer; } x_{ilsr}, x_{ilsd}^h, x_{ilsd}^{hc}, z_{ilksr}, z_{ilksd}^h, z_{ilksd}^{hc} \text{ are binary.} \quad (31)$$

The objective function (5) is the total cost, including fixed and variable costs in the regions, costs of renting or returning containers on the shipping day in the regions, repacking costs in the hub, fixed and variable costs for the pre-used containers and new containers in the hub, costs for renting or returning containers in the hub, and the penalty cost of cargo delay. Constraints (6)-(8) are the fixed and variable costs for containers in the regions, pre-used containers in the hub and new containers in the hub, respectively. The container quantity constraints in regions and hub are (9) and (10). Constraint (11) means that for each type of container, the quantity of using pre-used containers in the hub cannot be greater than the sum of containers used in all regions. Constraint (12) ensures that the number of each type of container used in the hub cannot exceed the limit. Constraint (13) makes sure that the pre-used container using plan should be the same in any scenario. Constraints (14)-(22) are container variable cost constraints. Constraints (23)-(24) are model robustness requirement, which means cargo delay is allowed. Constraints (25)-(31) are boundary restrictions.

Results and Analysis

A Case Study

A logistics company in Hong Kong provides air transport services worldwide. It collects shipping information from its customers including the characteristics for different types of cargoes, delivery dates, destinations and uncertain demand. The air cargoes need to be transported from two regions, Mainland China (Region A) and Vietnam (Region B), to the hub in Hong Kong first. The cargoes are unloaded and consolidated in Hong Kong before they are sent to two destinations, the EU (Destination α) and Northern America (Destination β). There are three types of cargo ($n = 3$): large, medium and small, with volume 1500, 1200 and 1000 cubic decimetres and weight 750, 600 and 500 kilograms, respectively. The company contacts an airline to rent air containers in advance. There are seven types of container ($m = 7$) for rent, and there is only one of each type of container available for rental in each region and hub ($L_{ir} = L_i^h = 1$ for each i and r). The airline provides the container rental information in Table 1, including the fixed cost ($), the volume limit (dm^3), the weight limit (kg), the breaking points (kg) and the unit charge rate ($/kg). The containers in the regions and hub have the same characteristics. If the company decides to continue using the containers in the hub, which just come from the regions, it will get 5% discount for the fixed cost, which means $\theta = 95\%$.

Table 1. Air container characteristics

Container type		1	2	3	4	5	6	7
Fixed cost		161617	105898	85207	74373	48713	46553	20695
Volume limit		6489	6300	5008	4882	3700	3150	1400
Weight limit		6800	5400	4200	4000	3900	3500	1200
Breaking point	a_{i1}	3968	2600	2092	1826	1196	1643	505
	a_{i2}	4722	3050	2490	2173	1423	1747	602
	a_{i3}	5290	3467	2789	2434	1594	2000	674
	a_{i4}	5976	3954	3149	2741	1825	2500	758

	a_{i5}	6273	4111	3307	2886	1917	2591	799
	a_{i6}	6800	5400	4200	4000	3900	3500	1200
Charged rate	δ_{i1}	0	0	0	0	0	0	0
	δ_{i2}	32	32	32	32	32	32	32
	δ_{i3}	0	0	0	0	0	0	0
	δ_{i4}	29	29	29	29	29	29	29
	δ_{i5}	0	0	0	0	0	0	0
	δ_{i6}	25	25	25	25	25	25	25

The uncertainty of cargo quantities of each type can be described by three scenarios: high demand s_1, medium demand s_2 and low demand s_3. Table 7 lists the cargo quantities under different scenarios. Table 3 shows the unit cost for returning unused containers and renting additional containers on the day of shipping. Those costs in the hub will be the same in all the regions. Table 3 also provides the unloading cost for each type of container in the hub.

Table 7. Cargo quantities under different scenarios

Scenario		s_1				s_2				s_3			
Region		A		B		A		B		A		B	
Destination		α	β	α	β	α	β	α	β	α	β	α	β
Cargo type	Large	2	2	2	3	2	1	1	1	1	1	1	1
	Medium	3	3	2	2	2	3	2	2	2	2	1	1
	Small	3	2	2	2	2	2	1	2	2	1	1	2

Table Error! No text of specified style in document.**8.** The unit cost of requiring/returning containers and unloading cost ($)

Container type	Unit cost of returning unused containers	Unit cost for requiring additional containers	Unloading cost
1	100000	200000	16000
2	70000	150000	10000
3	60000	120000	8000
4	50000	100000	7000
5	40000	80000	5000
6	35000	70000	4000
7	30000	60000	2000

Results

We use AIMMS 3.14 (with CPLEX 12.6 Solver) to solve the model. AIMMS 3.14 is an optimization software, which can be used to solve linear programming models. The model contains 3722 constraints and 2982 variables including 1960 integer variables. Table 4 and 5 show the results in the situation where the probabilities of high, medium and low demand scenarios are 80%, 10% and 10%. We let: $\omega_1 = \omega_1^h = 32000$, $\omega_2 = \omega_2^h = 28000$, $\omega_3 = \omega_3^h = 24000$, which represent the unit penalty costs for large, medium and small cargo in the regions and hub. Table Error! No text of specified style in document. presents the booking plan. Container 4, 5, 6 and 7 are booked in both in regions A and B because these containers are more economical than other types of container. One Container 4 and two Container 7 will be returned to regions when they arrive in the hub. There are no urgent returns or rentals occurring, due to the reduction of the variability cost. Table 5 gives the details of unshipped cargoes which means they will be considered in the following week.

Table Error! No text of specified style in document.Container booking plan

Container type		1	2	3	4	5	6	7
Region	A				1	1	1	1
	B				1	1	1	1
Pre-used container					1	2	2	
Hub						1	1	

Table 5. Unshipped cargoes

Cargo		Scenario s_1			Scenario s_2			Scenario s_3		
		L	M	S	L	M	S	L	M	S
Region	A	2	1				2			
	B	2		1						
Hub										

Conclusions

In this paper we build up a decision support system for the air cargo forwarding problems under uncertainty. Cargos need to be transported from regions to destinations via a hub. The air forwarders not only have to make a decision about the number of containers to be booked for the regions and hub in advance, before accurate customers' information becomes available, but also have to decide the number of extra containers to be required or the containers to be returned after the realization of uncertainty. We develop a robust optimization model to handle the uncertainty and flexibility, in which cargo delay is allowed. Computational results show that the robust model can provide a responsive and flexible decision support system to handle the uncertainty.

References

Yan, S., Lo, C.-T. and Shih, Y. L.: Cargo container loading plan model and solution method for international air express carriers. Transportation Planning and Technology, 29 (6), 445-470 (2006).

Wong, W.H., Zhang, A., Van Hui, Y. and Leung, L.C.: Optimal baggage-limit policy: airline passenger and cargo allocation. Transportation Science, 43 (3), 355-369 (2009).

Bing, L. and Bhatnagar, R.: Optimal capacity booking in air cargo transportation. Journal of Revenue & Pricing Management, 12 (3), 235-253 (2013).

Wu, Y.: Modelling containerisation of air cargo forwarding problems. *Production Planning and Control*, 19 (1), 2-11 (2008).

Wu, Y.: A dual-response forwarding approach for containerizing air cargoes under uncertainty, based on stochastic mixed 0-1 programming. European Journal of Operational Research, 207 (1), 152-164 (2010).

Wu, Y.: Modelling of containerized air cargo forwarding problems under uncertainty. Journal of the Operational Research Society, 62 (7), 1211-1226 (2011).

Posters

A systemic perspective on Racism in Football: the experience of the BRISWA project

S. Armenia, C. Carlini, J. Papathanasiou, A. Pompei, G. Tsaples
s.armenia@unilink.it, c.carlini@unilink.it, jasonp@uom.edu.gr, ales.pompei@gmail.com, gtsaples@gmail.com
web-page: http://www.briswa.eu

Abstract: To determine the important aspects of racism in football with a Systems Thinking approach, a series of workshops/sessions has been organized with experts in the field. The goal was to design a tool that could be used to get better insights into how racism in football emerges and where are the potential areas where policy makers could use as leverage for effective countermeasures. The diagram demonstrated the multidimensional nature of racism and the elements that could serve as leverage in potential mitigation policies. Some of the most interesting results include the following: the power structures of society and football should adapt to represent the actual demographic make-up of each country. Furthermore, policy makers should involve media more directly in every attempt to fight racism. Finally, racism in football is a mirror of racism in society. Hence any attempt to combat racism in football should be interlinked with corresponding efforts to fight discrimination in society.

INTRODUCTION

Racism in football is not a new phenomenon. There have been periods where it seemed to have vanished, through efforts to prohibit racist behaviour from and amongst the fans and governmental attempts to prevent racism through legislation. However, evidence suggest it remained.

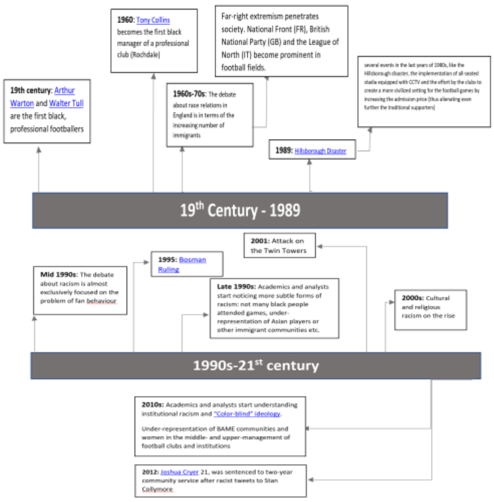


Figure 1: Evolution of Racism in Football (Timeline)

OBJECTIVES

The purpose of this paper is to report on the results from applying Systems Thinking, in the context of the BRISWA Project, in studying the phenomenon of racism in football:

- Understand the system of Football and Racism
- Gain insights into how the structure of the system drives racism
- Gain Insights into how potential countermeasures could mitigate the phenomenon or cause unwanted consequences

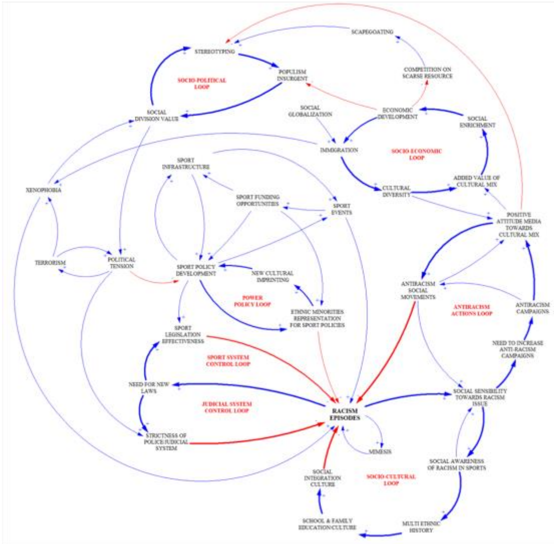
METHODS

Systems Thinking is a way of looking at systems from a holistic point of view:

- Its purpose is to determine what is the system's structure
- How the structure can affect its behavior over time
- It can be used in purely qualitative, quantitative fashion or a combination of the two
- Feedback loops, nonlinearities, stocks and flows

Group Model Building: A series of workshops with experts in the field

RESULTS



CONCLUSIONS

- Different societal, political and economic aspects could act on and affect racism in football.
- Fighting racism in football could not succeed solely with anti-racism campaigns focusing on fans' and players' behavior
- The power structures of society and football should embrace the multi-cultural character of modern society if they are truly to fight racism
- The judicial system of football should adapt to act more quickly when there is the suspicion that racism in football is on the rise
- the authorities should engage the media in their effort to combat racism

ANALYSIS OF THE USE OF ADDITIVE MANUFACTURING IN THE COMPONENT MANUFACTURING PROCESS

Giorgia Morgana , Sanguinetti Ricardo
Universidade Federal de Pernambuco - UFPE/ CTG Departamento de Engenharia Mecânica
Av. da Arquitetura, s/n Cidade Universitária CEP: 50.740-550 Recife-PE - *morgana_giorgia@Hotmail.com, ras@ufpe.br*

Abstract: Although the additive manufacture is a true story in Brazilian industries, its use in the mechanical industry segment has still been timid. The process used to manufacture parts / components used in the modeling of software by the decomposition of the material are still not practicable by companies, because they do not know the processes, equipment to be used and the benefits that the AM approach could offer. And with the advent of the financial crisis in the country, a technology that gives companies greater productivity under lower cost makes the competitiveness possible and, consequently, improves parts / components manufacturing processes. Among the benefits pointed out by several authors about the use of AM is: optimization of time and investment in launching new parts; waste elimination such as raw material; adaptable and easy technology; equipment with affordable costs and capable of producing complex materials; access to equipment / print parts in short term. Thus, this work, based on the analysis to be carried out, is of great relevance, because data will be collected and they will enhance the use of the MA, based on the several aspects that are considerable gaps in the literature associated with this subject, among which are: selection of components; and, processing and control of components.

INTRODUCTION

AM involves rapid prototyping, fast tooling and rapid manufacturing technologies, explains Volpato & Costa (2013). Since the evolution of new procedures, the fast prototyping process has been updated for the manufacture of components from a digital file (ALBERTI et al., 2014).
In this context, it has been possible to perform several materials processing by specific properties at specific component locations, as in the case of ceramic materials. According to (Gibson et al., 2010), this procedure in different materials was known as Multiple Materials Additive Manufacturing (MMAM). With that being said, it can be seen that in the engineering of materials and manufacturing, the development of new practices from the innovative process has been a trend in the market and of great relevance since, among some advantages, is the reduction to the losses in the components manufacturing process.
Thus, this project aims to study the impact in the manufacture of materials from the use of additive manufacture and with its multiple materials (polymers, ceramics and metals), in order to reduce the use of resources and with increasingly accessible costs. From this analysis, a structural model based on AM will be designed to support the search for efficiency in the manufacture of these components based on chemical materials / composition.

OBJECTIVES

This project aims to study the impact on the manufacture of components from the use of additive manufacture and its multiple materials, related to the structure of the materials layer, as in the case of polymers, ceramics and metals. For this purpose, the facts will be checked:

- Understanding from the types of materials how the processes in the additive manufacture work;
- Know the molds of materials manufacturing;
- Build hypotheses about the components;
- Adopt property sequence indicators;
- Analyze the manufacturing time of the components; Perform selection of process and materials;
- To propose a structural model applied to the *multicriteria* method, and so, use the AM approach efficiently.

METHODS

The research methodology adopted is based on a conceptual preparation based on bibliographic research on additive manufacturing, component selection and component processing and also control.
In addition to this procedure, an experimental study and application of a multicriteria method will be carried out, where the research steps begin with the exact formulation of the problem and the hypotheses, which delimit the precise and controlled variables that work on the problem studied and ordering the AM method (TRIVIÑOS, 1987).

RESULTS

The expected results are:

- Provide conceptual basis on the issues related to the project;
- To analyze the impact in the manufacture of components from the use of the additive manufacture and its multiple materials from the hypotheses raised;
- To propose a structural model based on the analyzes performed, and thus, to use the AM approach efficiently;
- Preparation and submission of paper for journals of relevance in the area.

Table 1: Additive Manufacturing Process

Process	Description	AM Method	Material
Fotopolimerização	A photocurable polymer is selectively cured using a light source Material is selectively deposited through an extruder head	Laser	Polymers
Modeling by extrusion		Heatting by a electrical resistence	Polymers, ceramics e metals
Pre deposited Fusion	An electron beam selectively regions a bed with pre-deposited powder	Laser e electron beam	Polymers, ceramics and metals
Cladding 3D	The addition material in powder form is injected directly into the bundle / pool	Laser e PTA	Metals and Ceramics

CONCLUSIONS

It was possible to analyze the impact in the manufacture of materials from the use of the additive manufacture and with its multiple materials (polymers, ceramics and metals), in order to reduce the use of resources and with increasingly accessible costs.

A Sequential Procedure for Improving the PCCM's Compatibility

María Teresa Escobar, Juan Aguarón, José María Moreno-Jiménez, Alberto Turón
Grupo Decisión Multicriterio Zaragoza (GDMZ), Universidad de Zaragoza, Spain
mescobar@unizar.es, aguaron@unizar.es, moreno@unizar.es, turon@unizar.es
web-page: http://gdmz.unizar.es

Abstract:

Consistency and compatibility are two topics usually employed in AHP-GDM. Based on consistency, a collective matrix, the Precise Consistency Consensus Matrix (PCCM), is constructed. This paper presents a sequential procedure for improving the compatibility of the PCCM. A case study with 5 alternatives and 3 decision makers illustrates the procedure.

CONTEXT / INTRODUCTION

- **Local** (one criterion) **AHP – GDM context**
  - *n* alternatives: *Ai* (*i* = 1,...,*n*)
  - *r* decision-makers: *Dk* (*k* = 1,...,*r*)
  - π_k : relative importance of *Dk* ($\pi_k \geq 0$; $\sum_{k=1}^{r} \pi_k = 1$)
  - $A^{(k)} = (a_{ij}^{(k)})$: pairwise comparison matrix of *Dk*
  - $A^{(G)} = (a_{ij}^{(G)})$: pairwise comparison matrix of the group

- **Consistency:** Internal coherence of the judgements in a pairwise comparison matrix
  - **Individual**

$$GCI^{(k)} = \frac{2}{(n-1)(n-2)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \log^2 \left(a_{ij}^{(k)} \frac{w_j^{(k)}}{w_i^{(k)}}\right)$$

  - **Collective**

$$GCI^{(G)} = \frac{2}{(n-1)(n-2)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \log^2 \left(a_{ij}^{(G)} \frac{w_j^{(G)}}{w_i^{(G)}}\right)$$

- **Compatibility:** Distance between the individual and collective positions
  - **Individual**

$$GCOMPI^{(k,G)} = \frac{2}{(n-1)(n-2)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \log^2 \left(a_{ij}^{(k)} \frac{w_j^{(G)}}{w_i^{(G)}}\right)$$

  - **Collective**

$$GCOMPI^{(G)} = \sum_{k=1,\dots,r} \pi_k \, GCOMPI^{(k,G)}$$

- **Use of the Consistency in AHP-GDM (Escobar et al., 2015)**
  - **Consistency Consensus Matrix** (CCM):
    - Interval judgement matrix that guarantees that all decision makers would simultaneously be consistent in their initial matrices
    - Limitation: sometimes incomplete
  - **Precise Consistency Consensus Matrix** (PCCM):
    - Precise judgment matrix (improvement of the CCM)
    - Good behaviour in consistency
    - Compatibility: can be improved

OBJECTIVES

- **Use** a **consensus matrix** based on the consistency (individual and collective): PCCM
- **Improve** the **compatibility** of the PCCM
  - Sequential procedure: one judgment per iteration
  - Selecting the judgements that most contribute to the overall compatibility of the group
  - Modifying the judgements using a combination of the initial value of the PCCM and the ratio of the priorities obtained with the AIJ procedure

REFERENCE

- Escobar, M.T.; Aguarón, J.; Moreno-Jiménez, J.M. (2015): Some extensions of the Precise Consistency Consensus Matrix. *Decis. Support Syst.*, 74, 67–77.

METHODOLOGY

Algorithm

- $\underline{a}_{ij}, \overline{a}_{ij}$ ($i, j = 1, \dots, n$): limits of the intervals of the Consistency Interval Judgement Matrix for the group (CIJM)
- $\theta \in [0,1]$
- $w^{(G)}$: priority vector obtained applying the RGM to the AIJ matrix
- *P*: a judgement matrix
- *V*: priority vector derived from *P* using the RGM method

Step 0: *Initialisation*

Let $t = 0$, $G^{(0)} = G$; $J = \{(i, j), \text{with } i < j\}$ and calculate for all $(i, j) \in J$:

$$d_{ij} = \sum_{k} \pi_k \log^2 \frac{a_{ij}^{(k)} v_j}{v_i}$$

Step 1: *Selection of the judgement*

Let (*r*, *s*) be the entry for which $d_{rs} = \max_{(i,j) \in J} d_{ij}$: $J = J - \{(r, s)\}$

Step 2: *Obtaining a PCCM entry*

$$G^{(t+1)} = G^{(t)}$$

$$\text{Let } z = \left(p_{rs}^{(t)}\right)^{\theta} \left(\frac{w_r^{(G)}}{w_s^{(G)}}\right)^{1-\theta}$$

$$p_{rs}^{(t+1)} = \begin{cases} \underline{a}_{rs} & \text{if } z < \underline{a}_{rs} \\ z & \text{if } \underline{a}_{rs} \leq z \leq \overline{a}_{rs} \\ \overline{a}_{rs} & \text{if } z > \overline{a}_{rs} \end{cases}$$

Step 3: *Finalisation*

If $J = \varnothing$, then Stop

Else let $t = t + 1$ and go to Step 1.

CASE STUDY - RESULTS

Real-life, public investment project for the restoration of the historical and cultural heritage of a village in Spain (Escobar et al., 2015)

- 5 alternatives: *Direct Use Value* (DUV); *Indirect Use Value* (IUV); *Potential Use Value* (PUV); *Existence Value* (EV) and *Bequest Value* (BV).
- 3 decision-makers (political parties) with weights: $\pi_1 = 5$; $\pi_2 = 4$; and $\pi_3 = 2$.

Table 1: Pairwise comparison matrices for the three political parties

PSOE	DUV	IUV	PUV	EV	BV
DUV	1	3	5	8	6
IUV		1	3	5	4
PUV			1	3	2
EV				1	1/3
BV					1

PP	DUV	IUV	PUV	EV	BV
DUV	1	3	7	9	5
IUV		1	3	7	1
PUV			1	5	1/5
EV				1	1/5
BV					1

PAR	DUV	IUV	PUV	EV	BV
DUV	1	5	7	7	5
IUV		1	1	5	1
PUV			1	5	1/3
EV				1	1/5
BV					1

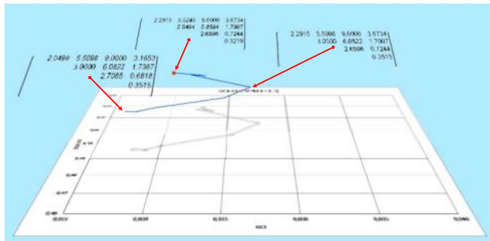


Figure 1: GCOMPI and GCI values for the sequential procedure (□=0.5)

CONCLUSIONS

Using the sequential procedure proposed:

- Compatibility is substantially improved
- Consistency is also improved

Future research

- Establish other criteria to determine the sequence of judgements to be modified

SUSTAIN GAME-BASED LEARNING ON URBAN SUSTAINABILITY

Co-funded by the Erasmus+ Programme of the European Union

ΙΔΡΥΜΑ ΚΡΑΤΙΚΩΝ ΥΠΟΤΡΟΦΙΩΝ IKY

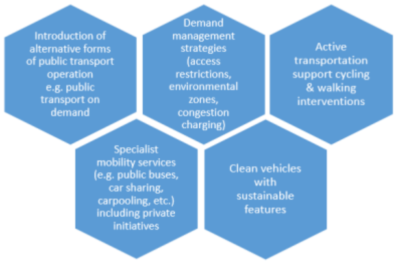
A project about Game Based Learning on Urban Sustainability

Jason Papathanasiou, Stefano Armenia, Federico Barnabe, Camillo Carlini, Natalia Ciobanu, Panagiota Digkoglou, Myrovali Glykeria, Lukasz Jarzabek, Maria Natalia Konstantinidou, Michalina Kulakowska, Alessandro Lanzuisi, Maria Morfoulaki, Camelia Onofrei, Alessandro Pompei, George Tsaples, Loukas Tsironis

Abstract: SUSTAIN is an ERASMUS+ project with an innovative perspective on urban transportation. Its target is to promote the importance of sustainability on the everyday problem of urban transportation among the students of higher education, which are the policy makers of tomorrow. In order to achieve its goals, the research team will develop a course that will be based on an interactive game with an analytical style of education. This game will allow students to learn about transportation sustainability and societal metabolism through playing. In addition, the research team will develop small and illustrative simulation models, which will make the definitions more concrete and allow students to experiment in a consequence-free environment.

INTRODUCTION

Traffic issues in cities create considerable problems, and as the years are passing by and the population grows in every big city, the number of vehicles also increases. Thus, creating a plan for sustainable solutions to the traffic problems of cities is essential.
The SUSTAIN project will create a game, which will be based on a process that will help every player to make choices for a sustainable urban plan. That game will be the base for a course for students. It is a quite innovative and hybrid perspective way of learning, in the sense that it will combine game-based learning with a cognitive and analytical style of education. It is essential to provide an innovative pedagogy to students of higher education, as they are the ones that will shape the future.



Sustainable mobility interventions

OBJECTIVES

Combining game-based learning with an analytical style of education, SUSTAIN aims to;

- Create small, illustrative simulation models that will make the definitions more concrete and allow students to experiment in a consequence-free environment with scenario exemplars.
- Create a Serious Game that will allow students to learn about transportation sustainability and societal metabolism through playing.
- Create a course dealing with transportation sustainability, societal metabolism and decision making under those contexts, while its techniques will be translated in everyday life, and formalize the mathematics necessary to make robust decisions. The course will be based on the Board Game.

METHODS

Sustain project is oriented on two main directions;

- **urban transportation sustainability** (measures and solutions able to tackle mobility challenges that lead to uncontrolled growth of vehicles and downgrading of quality of life) and
- **societal metabolism** (models that facilitate the description of flows of material and energy within cities and provide a framework to study the interaction between human and natural systems).

RESULTS

The results of the project are still at an early stage, but there are three Intellectual Outputs that are complete;

- **Output 1 – Ebook on societal metabolism.** It describes the main definitions of societal metabolism and the formal theoretical models, their advantages and disadvantages and how the theoretical notions of flows of energy and material can be translated to elements of everyday life.
- **Output 2 – Ebook on transportation sustainability.** It contains the definitions of sustainable transportation, the theoretical models that describe its behavior and discuss on the state of the art from a theoretical point of view on the research conducted on the issue.
- **Output 3 – Ebook on decision making in the context of sustainability.** Provides the formal mathematics and foundations of decision making especially in the context of sustainability, in order to students understand how the theoretical mathematics of decision making are translated into actual decisions. Linear programming, DEA, PROMETHEE and AHP are being detailed explained.

CONCLUSIONS

The SUSTAIN Project will deliver three additional Intellectual Outpus, one for simulation models, one for the board game that will translate the simulation models of O4 to game elements, mechanics and potential playing scenarios, and one for the board game itself.
Taking everything into consideration, SUSTAIN's team will try to boost its sustainable way of thinking, providing solutions through an interactive way of learning. Our goal is to promote the importance of a sustainable future for the urban transportation.

University of Macedonia (Coordinator) | System Dynamics Italian Chapter | CERTH/HIT | Centre for Systems Solutions | Ergo Ludo Editions | Society for Responsible Consumption

www.sustainerasmus.eu
sustain.erasmus@gmail.com

Find us on social media: SUSTAIN Project @SUSTAIN_Erasmus SUSTAIN-project-erasmus

APPENDIX

DECISION SUPPORT SYSTEMS - Historical Innovations and Modern Technology Challenges

Christer Carlsson and Pirkko Walden

Institute for Advanced Management Systems Research & Abo Akademi University, Auriga Business Center, 20100 Turku, Finland
christer.carlsson@abo.fi, pirkko.walden@abo.fi

Abstract. Managerial tasks carry latent needs for *support to do a better job*; classical DSS had at its core the approach *to support, not replace*. We worked out a DSS called *Woodstrat* for strategic planning and management and could verify – in full-scale implementation - the DSS characteristics Sprague worked out and most of the DSS benefits Keen had found. This brings out the historical innovations that DSS introduced. We also found that a relevant and useful DSS could help "self-confident professionals" to back away from predictions on future demand and competition that could not find support in facts. The digital disruption of the 2010'es brought big data and the need for decision making in almost real time. It also introduced analytics and faster, more effective algorithms developed in computational intelligence. The ICT of the 2010'es invites decision-making anywhere and anytime which is a challenge for the cognitive ability of decision makers. The road map for DSS for the 2020'es points to digital coaching systems that adapt to the cognitive levels of the users.

Keywords: Classical DSS, modern DSS, computational intelligence.

1 Introduction

The ICDSST 2019 offers a possibility to look at 40 years of DSS history, the innovations that DSS introduced and the successes these innovations brought. There has also been expectations on successes and breakthroughs that did not happen, and DSS promoters were too enthusiastic in some cases in promising improvements in productivity and profitability. Nevertheless, the DSS brought some innovations that have stood the test of time and have returned – repeatedly - only changing shape to new forms as DSS technology has developed.

We have some history ourselves with DSS – the first IFPS-based DSS projects were started in Finland in 1985 and a research group that later formed the Institute for Advanced Management Systems Research (IAMSR) was the driving force. The first DSS applications built on the original Fortran-based IFPS mainframe version and were challenging to design and build. Mainframes were inflexible to run as decision support tools and it required creativity to get models, algorithms and solutions close enough to real world problem solving to get valid and relevant support to the users (cf. [2], [3], [25]).

The first DSS conference was the DSS-81 in Atlanta, Georgia sponsored by Execuom Systems Corporation, the developer of the IFPS software system. Gerald R.

Wagner, then President of Execucom, and Peter G. W. Keen initiated this conference and more or less defined the agenda for the development and use of DSS. Peter Keen had a keynote address – "*Decision Support Systems – Lessons for the 80's*" – and Ralph H. Sprague gave a tutorial on *Decision Support Systems* in which he defined all the key elements of a viable DSS. Several papers pointed out technology challenges that then triggered development efforts for successive versions of the IFPS (and competing software).

Most of the early authors note the Gorry and Scott Morton [12] paper ("*A Framework for Management Information Systems*") in Sloan Management Review in 1971 as the starting point for DSS. The paper builds on Scott Morton's doctoral thesis [20] at Harvard Business School in 1971 that outlined "management decision systems". The Sprague [23] paper ("*Framework for the Development of Decision Support Systems*") in MIS Quarterly in 1980 then summarized all the essential elements for the design, development, implementation and use of decision support systems.

The keynote of Peter Keen and the tutorial of Ralph Sprague at DSS-81 summarize the key innovations that decided the emergence and the success of decision support systems. We need to note the context in the early 1980'es: Data Processing (DP) dominated how managers and corporations viewed the use of computers for management. The focus was on cost effectiveness and productivity, the systems were mainly run on mainframe computers (IBM was an overwhelming market leader), the dominating software were Cobol and Fortran, information systems were large, complex and inflexible, and investment costs were high. The agenda presented by the DSS pioneers did not in many cases get friendly responses; in some cases, reactions were outright hostile [16].

DSS builders focused on the users' priorities, they developed systems linked to key business activities and they viewed the quality of a system from the value it gives to the users rather than the level of (advanced) technology applied.

DSS reflects demand economics: service, fast delivery, ease of use, benefit focused more than cost, imprecision allowed for timely delivery and user control.

Early case studies by Keen [15] showed a number of benefits identified by DSS users: (i) increase in the number of alternatives examined; (ii) better understanding of the business; (iii) fast response to unexpected situations; (iv) ability to carry out ad hoc analysis; (v) new insights and learning; (vi) improved communication; (vii) improved management control; (viii) cost savings; (ix) better decisions; (x) more effective team work; (xi) time savings; (xii) making better use of data. These and similar benefits still appear in the literature, even if the underlying DSS technology has changed several times and the technology gets a different label than DSS.

Managerial tasks are not routine and the latent needs they create are for "*support to do a better job*", which is an informal DSS credo.

The DSS architecture builds on mainly three components: (i) a dialogue manager/interface between the user and functional routines; (ii) a data manager; (iii) functional routines. This approach is a distinctive technology contribution of DSS.

The philosophical, attitudinal core of Decision Support is "*support, not replace*". It is impossible to support individuals if we do not know what they do, how they think, what doing a "*better job*" means to them, and what they need to have (cf. Keen [14], p.

190). In the 1970's and 1980's the prevailing management science paradigm [26] developed a "black box" approach to better decisions. In case human cognitive ability was not enough optimization algorithms took over (replaced, if we like) and offered the best possible solution. The algorithms were quite often beyond the knowledge and skills of the users who sometimes did not see why optimal solutions would be the best possible in any given problem situation. The DSS addressed this problem and promoted problem solving that built on managers' intuitive understanding and experience of how to solve problems. The reasoning was simple – there is no need to "*sell*" solutions if managers (the problem owners) run the problem solving process with some support from computer based technology.

Sprague [24] found it more useful to collect the "characteristics" of DSS than to try formal definitions or to distil some common understanding from actual use cases. He collected the following "characteristics" from several authors ([24], 193-194): (i) DSS aim at the less well-structured, underspecified problems of upper level management; (ii) DSS combine the use of models or analytic techniques with traditional data access and retrieval functions; (iii) DSS focus on features which make them easy to use by non-computer people in an interactive mode, and (iv) DSS emphasize flexibility and adaptability to accommodate changes in the environment and the decision making approach of the user.

A distinctive feature of the early descriptions of DSS is that it should support all phases of decision-making. Sprague [23] connects this to Simon [22]: intelligence (environment search for decision needs), design (inventing, developing and analyzing action alternatives) and choice (selecting a particular action from available alternatives).

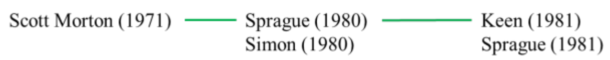


Fig.1. Intellectual foundations for DSS

A final distinguishing feature of DSS is the iterative design [24]. The typical four steps of an information systems development process – analysis, design, construction, and implementation – combine in a single step, which repeats iteratively. A typical process starts with the manager and the DSS builder agreeing on a small but significant sub-problem, designing support functions needed for decision-making and collecting experience of the functionality. Then another sub-problem is tackled with the same approach and when the decisions are sufficiently good, the two DSS modules are connected to allow the solutions to be integrated. Then the process continues over *n* subproblems and *m* modules until we have a full DSS construct. The integration of subproblem solutions tends to offer challenges [3].

The rest of the paper will address the key issues of DSS from two perspectives: experience gained from actual work with DSS to test the benefits and characteristics of DSS that Keen and Sprague outlined. Secondly, we will find out if the core ideas of DSS are still relevant with the technology and the decision-making contexts of the 2010's and 2020's. Section *two* works through some of the controversies with introducing computer support for strategic management. In section *three*, we will work through experience gained from a DSS developed and used for strategic planning and management. Section *four*, works out some key principles on decision-making. In section *five*, we will introduce analytics and the requirements of a digital economy. Section *six*, outlines some promising design principles for the DSS of the 2020's.

2 Computer Support for Strategic Management

The context we have is work we carried out with 11 strategic business units (SBU) in a forest industry corporation to help them make their annual strategic planning process both more detailed and faster with computer support. The timeline was the mid-1990's and the support technology we used is now outdated. Even the strategic business units have merged, split and reorganized several times. The corporation has adapted to changing customer needs and markets, to new and advanced production technology and to new competitors. Nevertheless, we have found that the experience we gained offers a useful illustration of the DSS visions we collected from the early pioneers and the *Transactions of DSS-81*. It turned out, that even with the rather rudimentary technology we applied (compared with the possibilities now offered) we could support strategic decisions that SBU managers made for their real-world operations.

At the time, there was some debate about strategic planning vs. strategic management (cf. Mintzberg [19]) and the SBU-managers wanted to form a joint understanding with us. Thus, we agreed that *strategic management* is the process through which a company for a chosen planning period (i) defines its operational context, (ii) outlines and decides upon its strategic goals and long-term objectives, (iii) explores and decides upon its strengths, weaknesses, opportunities and threats, (iv) formulates its sustainable competitive advantages and (v) develops a program of actions. The actions exploit its competitive advantages and ensure profitability, financial balance, adaptability to sudden changes and a sound development of its capital structure. This lengthy joint understanding changed a number of times until the SBU-managers agreed that it makes sense to them and their own SBUs. We learned that this is a crucial step for the development of support technology – unless the users have sufficient understanding of the process we are going to support the possibilities for success will be rather slim.

As the conceptual framework was in place the decision support technology should provide a platform to deal with practical issues. Sufficient and reliable data on markets and competitors needs first to be stored in usable form for the strategic planning (previously corporate planners sent out macro-level reports that were mostly irrelevant for the SBU's). SBU-managers had experience of strengths and weaknesses of key competitors but needed tools to work out their insight and build data for strategic planning. It turned out that SBU-managers also had good perceptions of their own competitive advantages relative to their competitors and their own competitive positions in key market segments. Again, they needed tools to work out their insight as data for strategic planning. The final step, connecting competitive and market positions with productivity and profitability for an SBU and then with a financial position and capital structure offered more challenges and the help from support tools was very welcome.

In the mid 1990's the Mintzberg ideal for strategic management stressed the notion of an *emerging strategy*, which (simply stated) built on continuous dialogue among senior managers about present and future markets, competitors and relative competitive positions that would decide strategic positions, return on assets and shareholder value (cf. [19]). The dialogue would converge to consensus on future directions through a viable conceptual framework, that Mintzberg offered and that would guide the managers to find a joint understanding of the emerging strategy.

Mintzberg did not believe in computer support for senior managers and he quite emphatically stated that computers have no place in strategic management [18]. We need to remember that the context for Mintzberg's position built on the (mainframe) computer technology of the 1970's and that senior managers did not operate computers at that time. Nevertheless, his conclusions on the role of computers – still widely quoted in the 1990's – was wrong. Eden [10] demonstrated with his *Decision Explorer* that computer support is very useful for cognitive mapping, that the software is manageable for (senior) managers and that computer support is instrumental for a Mintzberg dialogue on future directions.

3 The *Woodstrat* Decision Support System

In the work on *Woodstrat* we got inspiration from Eden's systems constructs and then applied some new principles for *hyperknowledge* that Chang, Holsapple and Whinston had introduced [7]. A decision takes form through navigation in a universe of concepts. Some of the concepts are descriptive, some are procedural, and some are context-dependent, abstract goal formulation and motivation concepts that serve as instruments to forge a joint value and goal system. The hyperknowledge process will interlink the concepts to allow the impact of changes in one concept to be worked out in another concept (cf. fig.2 –interlinking shown with blue, green and red lines). For the *Woodstrat* the interdependences represented the internal logic of an SBU business context.

Fig.2 shows the overall structure of the *Woodstrat* and the strategy formation process.

The *Woodstrat* took form through a series of prototypes (cf. the iterative DSS building approach). The first versions used a *LISP*-based expert system shell, which proved too inflexible for the internal logic of the business context. The next series of prototypes used *Toolbook* to introduce the hyperknowledge constructs. This platform was too hard to implement for managers who are not skilled software users. The full-scale system took form as a hybrid system in *Visual Basic* in which we rewrote the *LISP* and *Toolbook* constructs as objects. The *Visual Basic* offered graphical user interfaces, multiple-document interfaces, object linking and embedding, dynamic data exchange, effective graphics and custom controls with procedures from dynamic-link libraries. We built in *what* if- and *goal* seeking features that had proved very useful in IFPS. These features are now available – even if most users do not even realize it – in *Microsoft Excel*, in further developed and advanced forms.

The *Woodstrat* supported the strategy formation process of an SBU; we designed and built versions of the DSS for 11 SBU's with the help of 40 senior managers who worked on the annual strategic planning of the division. The Senior VP of the division was project champion, as he wanted to mobilize the experience, insight and intuitive strategic knowledge of his managers, something that he had found that the standard corporate planning process did not capture.

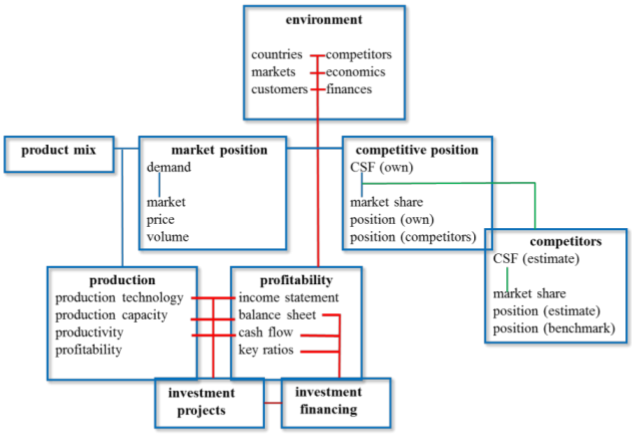


Fig.2. The *Woodstrat* decision support system for strategic management [3]

An SBU operates in several countries, with well-defined product groups and specified customer segments. Markets and segments differ for different product groups, and their importance varies over the planning period. The strategic market positions (MP) are determined hierarchically: segments are defined for each product group and product groups are selected for each country; for each segment demand and price development estimates are made and consolidated to product group and country levels. The weighted average of growth and price development estimates update the estimated *net sales* line in the income statement through functional links.

We built rather an extensive database of country specific economic indicators and related forecasts as part of *Woodstrat* to which we added market and segment specific forecasts on the development of price and demand levels. It turned out that this helped the SBU managers to anchor and calibrate their assessments of growth and price developments. This offered a base of facts for the strategic visions.

The competitive positions (CP) are activated with logical links from the same base of facts as the MP, and the MP and CP are worked out in parallel. The CP builds on critical success factors (CSF), which are SBU-specific and worked out in a series of seminars with the SBU-managers. The CSF are calibrated with relative changes to the previous year and changes to the CP. This process changes and adapts visions of the MP when evaluated against the CSF and the relative strength of the competitors.

Three selected competitors were evaluated on the same CSF in a benchmarking workshop with each SBU. Effort and time was spent to identify "good" competitors, i.e. competitors that had managed to take away good customers and have an impact on the MP. CSF and CP averages were determined for the competitors; the relative differences in CP were calculated and used to assess the relative strategic CP for the SBU. The CP estimates were linked to the MP and used to calculate an estimated development in volumes and prices; here we had built in a function for the SBU managers to override the estimates with their own estimates on volumes and prices. The principle was that the managers' active customer relationships should decide the MP.

We used graphs to summarize relative competitive positions, the expected market development and total sales. This again proved to be useful as the SBU managers wanted to get graphical overviews but had learned in previous years that the corporate planners could produce that for them "only with considerable difficulty".

The production position (PRO) estimates productivity as consequence of the MP and the CP. *Production sold* is determined and transferred from the growth and price development specified in the MP. *Productivity* is determined from several factors – labor, raw material, electricity, steam and technology. The module has functions for *profitability* and *capacity limits*. The cost factors of the income statement update the productivity factors with knowledge-based links. The productivity and profitability measures are numerical functions of the CSF and the visions implemented in the MP and CP.

There is an SBU Report activated from the summary level of the MP module with MP data and raw material costs. In the report, there is a projected income statement, linked with a balance sheet, a statement of funds and a report on key ratios. The modules update each other through knowledge-based links that follow proper accounting principles. The main key ratio followed by the forest industry is the *return on net assets* (RONA). We added a *what if* type of graphical RONA simulation to the report which allowed SBU managers to find critical sales or operating cost levels for reaching target levels of the RONA. A major benefit of the linked modules is that an SBU manager can work out several MP and CP scenarios and quickly find out how reasonable they are in terms of the division's RONA targets. The linked modules were also major time savers.

A *Woodstrat* feature the SBU managers much wanted allowed them to work with investment plans interlinked with financing plans and further linked to *net sales*, *cash flows* and *key ratios*. This allowed them to demonstrate and motivate the impact and consequences of the investments they wanted the corporation to accept and fund. The existing corporate policy routinely rejected investment proposals without any changes to revenues; now the SBU managers could demonstrate the impact on *net sales*, *cash flows* and *key ratios* if an investment proposal is not approved (to the surprise of corporate planners who were not used to fact-based arguments and negotiations).

We included a *Memo* module to allow the SBU managers to keep track of their assumptions, knowledge points and motivations for market and competitor estimates. A number of factors were not well known and registered in the *Memo* for follow-up studies; a number of questions and ideas went to sales offices in Europe for verification and collection of more and better data. We were able to collect and analyze the *Memo* material from the 40 SBU managers, which gave rather unique insight in the strategy formation process that the DSS guided and supported. Mintzberg [19] would probably not have agreed, but we noted that the DSS helped managers to formulate strategic visions for the business context.

We carried out systematic follow-up studies with the SBU-managers to find functions that needed improvement and links that at some point would produce invalid outcomes. We also collected some positive evaluations: "*the system guides the user to focus on important issues which eliminates unnecessary work*". Also, "*compared with my old way I worked more thoroughly and used more time than before*". On the DSS, "*the DSS captured us – the drawback was that we concentrated too much on details in the MP and CP*". Finally, "*the planning process became real team work*".

An evaluation of the *Woodstrat* experience shows that strategic planning and management fulfills (i) in Sprague's list of DSS "characteristics" [23], [24]. The support system works with interlinked modules (models with algorithms and hyperknowledge links) that use *Visual Basic* dynamic data exchange for data access and retrieval (cf. (ii)). The SBU managers worked interactively with us on the *Woodstrat* design, implementation and use, which resulted in functionality suited to non-computer people (cf. (iii)). The support system design aimed at an adaptive platform, that supports strategic planning for the next 3-5 years (cf. (iv)). The follow-up studies with the SBU-managers verified that the benefits that Keen [15] had collected exist also for *Woodstrat*. We identified (i)-(ii), (iv)-(vii), (ix)-(xii).

4 Support for Decision Making

We will now change context from the history of DSS to the 2010-2020's and the challenges of the growing digital economy. Decision support systems have decision-making at their core and we propose that this core will be the same also in the digital economy. Decision support has to tackle the fast growth of big data, which invites proposals that things will be more complex and difficult in the 2020's than in the 1970's. Streaming big data now appears to make algorithms and modelling impractical as the huge amounts of data will take too much time to process, which again will make decision-making too slow. Fast decision-making in almost real-time is a necessity in the digital economy ("the fast eat the slow" as the slogan goes). Kahneman [13] shows that fast decision-making in many/most cases will produce bad decisions; a good credo for the DSS to follow is – "if there is time to make bad decisions, there should be time to make good or better decisions".

Zeleny [26] wrote a classical contribution to decision making. First, with a single attribute or objective or utility function there is no decision-making involved, the decision is implicit in the measurement and becomes explicit in the search for a best value. With multiple criteria (attributes, objectives) or value functions, we get actual decision-making. As a human process – also when guided by DSS – decision-making is dynamic and composed of partial decisions in pre-decision, decision and post-decision stages. The three stages require support from different kinds of data sources, data, information (knowledge), modelling tools and experiments, which all should be part of the DSS constructs (still consistent with Sprague [24] and Simon [22]: intelligence, design, choice).

Kahneman [13] offers numerous examples of how limitations to human cognitive ability create bias when we want to address future uncertainties. In the *Woodstrat* cases, we had to build foresight to guide business decisions for 3-5 years into the future. SBU managers had to understand customers, markets, competitors and future economic and financial scenarios in order to find reasonable and valid estimates of demand, prices and market shares (the actual process was a bit more detailed and complex). Then managers face what Kahneman [13] calls *vivid outcomes*. Probability estimates of future outcomes are sensitive to how much detail we know and use. Probability estimates become too optimistic with positive details or too pessimistic with negative details. Probability estimates are subjective and may give very wrong impressions of the future. In contrast, DSS offers a factual database and tools for objective estimates. Many strategic planning scenarios turned out to be far off the mark in the SBU's before the *Woodstrat*.

Kahneman [13] raises a sensitive issue – "*when can you trust a self-confident professional who claims to have an intuition*". SBU managers are professionals; they have been working with their products, customers, competitors and markets for years. It is a difficult process to challenge their intuition on future development of key strategic factors. These include demand, possibly competitive prices, relative market and competitive positions, raw material and operative costs. They also include uncertain facts about future economic scenarios for the countries in which they operate. Kahneman [13] simply states that it is wrong to blame anybody for inaccurate forecasts in an unpredictable world. It turned out that *Woodstrat* helped the professionals to test, adjust and correct their initial intuitive forecasts without drama.

5 Decision Support for the 2020's

In a recent report called "*Competing in 2020: Winners and Losers in the Digital Economy*" [8] Harvard Business Review worked out the impact digitalization will have in a few key industrial sectors. The method was a multinational survey aimed at senior managers, executives and board members; 783 respondents completed the survey; all of them indicated that they are digital decision makers or influencers. The key industries covered were manufacturing or resources, financial services and technology, mainly organizations with more than 10 000 employees.

Among the respondents 16% stated that their companies are *digital* (most products/operations depend on digital technology), 23 % that they are *non-digital* (few if any products/operations depend on digital technology) and 61% that they are *hybrid* (some products/operations depend on digital technology).

The business world changes taking place are "*the digital disruption*" and "*the digital revolution*". The contention is that digitalization will have significant impact on both the structure and the operations of the business world, on the business models and on how companies cope with increasing competition, slimmer margins for productivity and profitability and growing requirements for effective planning, problem solving and decision-making. Digitalization is of course bringing opportunities: the two most significant are enhanced customer relationships that allows to work out (and charge for) individual value adding in ways that have not been possible before and value chain integration that offers control of markets and rapid market changes with much better tools.

The report found a significant performance gap between digital leaders ("*digitals*") and the rest (called "*non-digitals*"). It shows that 84% of the digitals use big data and analytics, but only 34% of the non-digitals; 51% of the digitals use cognitive computing/AI, but only 7% of the non-digitals. Another significant difference - the digitals have data science and data engineering on staff (62%), the non-digitals much fewer (20%); all professionals working for the digitals have the ability to work with and make sense of data and analytics (76%), not that common for the non-digitals [30%]. The conclusion is that a strong analytics capability is key to digital business – companies that want to compete in the digital economy will have to invest in analytics people, processes and technology. The message is – curiously enough – the same we learned from Keen [14] and Sprague [24] almost 40 years ago but the context (digitalization) and the modelling methods (analytics) are now very different; how different we will find out.

In their policy statement for the new *Journal of Business Analytics* Delen and Ram [9] show in a word cloud analysis (cf. fig.3) that big data – analytics – (text) mining over the last decade started to appear as related concepts in journals and conference publications. This is not surprising as digitalization produces fast growing sets of big data and it is now evident that analytics offers useful tools to cope with big data.

Fig.3. Analytics and Big Data [9]

Delen and Ram [9] also offer an overview of the evolution of (business) analytics that shows it as growing out of the DSS movement in the 1970's (cf. fig.4).

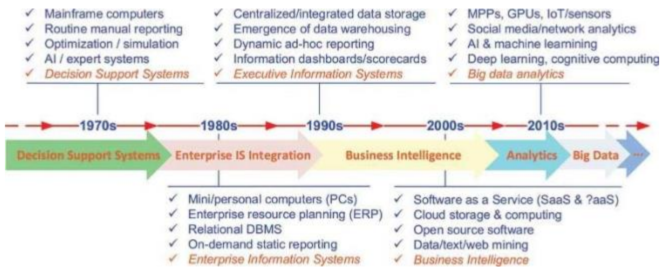


Fig.4. A historical view to the evolution of analytics terminology [9]

Business analytics has three functional orientations: descriptive, predictive and prescriptive; INFORMS has the same specification of analytics – descriptive ~ business intelligence, predictive & prescriptive ~ advanced analytics. DSS literature usually lists the functions specified [16] (cf. also [1]):

(i) *Descriptive* – business reporting, dashboards, scorecards, data warehousing
(ii) *Predictive* – data mining, text mining, web/media mining, machine learning
(iii) *Prescriptive* – optimization, simulation, algorithms, network science

The classifications are not precise and exhaustive, e.g. machine learning methods and simulation models appear in descriptive modelling.

Visions similar to Delen and Ram of the possible developments of DSS methods and applications appeared earlier, which anchors business analytics as part of the decades-long traditions. In *Past, Present, and Future of Decision Support Technology* [21], in a special issue of the *DSS journal* [4] published in 2002, the starting point was Keen's agenda for DSS entering the 1990's [16] and the aim was to work out the most promising research areas based on new technology. Keen (in 1987) wanted DSS developers to apply analytic models and methods for a more prescriptive view of how to make decisions that are more effective (he also wanted focus on "*decisions that matter*"). Keen encouraged DSS developers to exploit software tools and AI to make DSS move towards semi-expert systems. We followed up on Keen's proposals in [16] and worked out the following agenda:

(i) identify areas where tools can transform qualitative insight and uncertain and incomplete data into useful knowledge;
(ii) use intelligent systems and methods for prescriptive, more effective decision making;
(iii) exploit advanced software tools to improve the productivity of decision and working time, and,
(iv) assist and guide DSS practitioners for effective decision making

These guidelines are general and open-ended but they still make sense and are useful in the present business context of digitalization and big data. A key difference between 2002 and 2018 is the portfolio of tools we have available for building decision support.

The algorithms belong to the computational intelligence family, are faster, more powerful, and can handle (very) big data. The user support is adaptive and interactive, and it will evolve with the cognitive ability of the user. The platforms build on smartphones, tablets, laptops and powerful desktops to provide users with real-time decision support wherever they are and whenever they need it. The users, however, still need to provide the cognitive ability, the experience, the intelligence and the insight to make effective and better decisions.

6 Computational Intelligence in Decision Support

The digital economy and the big data challenges appear to disqualify the classical algorithmic methods (cf. [6]) – optimal problem solving is nice but useless if it cannot meet the hectic pace of the digital economy. Classical algorithms cannot process big data in reasonable time – or even not at all. Some of these claims are fallacies – it is not necessary to process big data at all if we first use classical statistical methods to find the smaller subsets of factors that are relevant and actually influence the problems we need to solve (cf. [6]). The classical algorithms are again relevant for the smaller subsets.

In case we actually have to work with big data it appears that we should look to *computational intelligence* algorithms, which offer to be much faster than polynomial methods [17]: neural networks, support vector machines, genetic algorithms, genetic programming, swarm intelligence, software agents and soft computing. There is a drawback, users need to have some fairly advanced mathematical and software skills to operate computational intelligence.

There is a central challenge in digitalization; the human users of advanced automated systems are the weak links (cf. [5]). Large, automated systems rely on advanced algorithms and large complex computational systems; it is not self-evident that human system users have the knowledge and/or the skills to manage the systems and to operate them to produce a competitive RONA.

System users have diverse backgrounds and different levels of experience. Some users understand everything and master the systems in a short time; then they will start to contribute to development. On the other hand, some users are slow to learn and/or are not motivated; it will take time for them to reach even minimal acceptable levels.

The D2I joint industry and university research program [5] proposed that we build on human and system joint intelligence for digitalization, that we use fast, automatic algorithms for large, well-structured datasets and combine this with knowledge mobilized from seasoned context experts. In order to make it work, human systems users need context relevant advice (in real time, with real data and information) that is adapted to their cognitive abilities and background knowledge (i.e., advice they can understand and use). *This could be the mission statement for the DSS of the 2020's.*

The *digital coaching systems* got started a few years ago (cf. [11]) as an answer to the demand on human operators to master advanced automated systems in complex and very large industrial process systems. Digital coaching will work with data that is collected from digital devices, instruments, tools, monitoring systems, sensor systems, software systems, data and knowledge bases, data warehouses, etc. and then processed to be usable for the digital systems that will guide and support users.

Digital coaching requires that we master the transition from data to information, and on to knowledge, also known as *digital fusion*. Data fusion collects and harmonizes data from a variety of sources with different formats and labels. Information fusion uses analytics to build syntheses of data to describe, explain and predict key features for problem solving and decision-making. Knowledge fusion uses ontology to build and formalize insight from data and information fusion as a basis for computational intelligence methods, AI, machine learning, soft computing, approximate reasoning, etc. The early versions of DSS hinted at the need for what we now call digital fusion (cf. [5], [6]) but which lacked the necessary software tools. They are now available and appear to be on a path towards becoming both more intelligent and effective.

The DSS of 2020'es will quite possibly be digital coaching systems that will guide users in the digital economy over smartphones, tablets, laptops, terminals to cloud services and new digital support devices that will appear as part of the environment.

References

1. Arnott, D. and Pervan, G. 2008. Eight key issues for the decision support systems discipline. Decision Support Systems (44), (2008), 657–67
2. Carlsson, C. and Walden, P.: Strategic Management with a Hyperknowledge Support System, Proceeding of the HICSS-27 Conference, (1994), 241-250
3. Carlsson, C. and Walden, P.: Cognitive Maps and a Hyperknowledge Support System in Strategic Management, Group Decision and Negotiation, (6), (1996), 7-36
4. Carlsson, C. and Turban, E.: DSS: Directions for the Next Decade, Decision Support Systems (33), (2002), 105–110
5. Carlsson, C.: Analytics Mobilized with Digital Coaching, Intelligent Systems in Accounting, Finance and Management, ISAF1421, January/March, 2018, pp 3-17
6. Carlsson, C.: Decision Analytics – Key to Digitalization, Information Sciences, (460-461), September 2018, 424-438
7. Chang, A.-M., Holsapple, C.W. and Whinston, A.B.: Model Management Issues and Directions, Decision Support Systems, (9), (1993), 19-37
8. Competing in 2020: Winners and Losers in the Digital Economy (2017), A Harvard Business Review Analytic Services Report, April 25, 2017
9. Delen, D. and Ram, S.: (2018) Research Challenges and Opportunities in Business Analytics, Journal of Business Analytics, (1)1, (2018), 2-12
10. Eden, C.: Strategy Development and Implementation – Cognitive Mapping for Group Support. In: Hendry, J. and Johnson, G. (eds.) Strategic Thinking: Leadership and the Management of Change, Wiley (1993)
11. Fern, A., Natarajan, S., Judah, K., and Tadepalli, P.: A Decision-Theoretic Model of Assistance, Journal of Artificial Intelligence Research, (49) 71-104, (2014)
12. Gorry, G.A. and Scott Morton, M.S. : A framework for Management Information Systems, Sloan Management Review, 13(1), 55-70 (1971)
13. Kahneman, D.: Thinking, Fast and Slow, Penguin Books (2011)
14. Keen, P.G.W.: Decision Support Systems – Lessons for the 80's. In: Young, D. and Keen, P.G.W. (eds.) DSS-81 Transactions, Atlanta, Georgia (1981), 187-192

15. Keen, P.G.W.: Information Systems and Organizational Change, Communications of the ACM, 24(1), 24-33 (1981)
16. Keen, P.G.W.: Decision Support Systems: The Next Decade, Decision Support Systems 3 (3), 253– 265 (1987)
17. Kordon, A.K.: Applying Computational Intelligence, Springer, Berlin (2010)
18. Mintzberg, H.: Patterns in Strategy Formation, Management Science, 24(9), (1978)
19. Mintzberg, H.: The Rise and Fall of Strategic Planning, Prentice-Hall (1994)
20. Scott Morton, M.S.: Management Decision Systems: Computer Based Support for Decision Making, Division of Research, Harvard University, Cambridge, Mass. (1971)
21. Shim, J. P., Warkentin, M., Courtney, J.F., Power, D.J., Sharda, R. and Carlsson, C.: Past, Present, and Future of Decision Support Technology. Decision Support Systems, (33) 2, (2002), 111–126.
22. Simon, H.A.: Cognitive Science: The Newest Science of the Artificial, Cognitive Science, 4, 33-46 (1980)
23. Sprague, R.H.: Framework for the Development of Decision Support Systems, MIS Quarterly, 4(4) (1980)
24. Sprague, R.H.: Decision Support Systems: A Tutorial. In: Young, D. and Keen, P.G.W. (eds.) DSS-81 Transactions, Atlanta, Georgia (1981), 193-203
25. Walden, P. and Carlsson, C.: More Effective Strategic Management with Hyperknowledge: Case Woodstrat. In Darzentas, J., Darzentas, J.S. and Spyrou, T.: Perspectives on DSS, University of the Aegean (1995), 139-156
26. Zeleny, M.: Multiple Criteria Decision Making, McGraw-Hill, New York 1982

30 years Business Intelligence : from Data Analytics to Big Data

Isabelle Linden[1]

NADI institute, FOCUS research group
University of Namur, Namur, Belgium
isabelle.linden@unamur.be,
https://directory.unamur.be/staff/ilinden/

Abstract. At the crossing of disciplines as Information Systems, Managment, Decision Support Systems, Data Mining, and Data Visualisation, Business Intelligence (BI) is understood in very different ways by the multiple concerned actors. This paper aims to offer to all of them an integrated view on multiple perspectives. To this end it first proposes a standard Business Intelligence approach. Then, it describes the main technical challenges addressed in the litterature with a particular focus on those risen by the emergence of Big Data.

Keywords: Business Intelligence, Big Data, Decision Support Systems

1 A brief history of Business Intelligence

Since men have been involved in production and trade activities, and probably more critically since the industrial revolution, there have always been people to analyse their performance and question their optimization. Statistics and later data mining offered powerful tools to support this type of quest.

The 1980s and 1990s saw the explosion of computerization in organizations. Many data and information previously processed by hand on paper were digitized. Digital data sources multiplied not only in administration but also at the very heart of production chains and processes.

At the same time, processors gained in power, memories in capacity and algorithms in efficiency. Such convergence has offered analysts unprecedented processing capabilities. They have extremely wide fields of exploration to investigate. But often, the dream becomes a nightmare when it comes to supporting top management and analysing issues that cut across their organization. Indeed, analyses require then access to information disseminated in various and varied systems. Moreover, data sources are not only multiple but also heterogeneous in their formats and structures.

Specific architectures and platforms emerged in the 1990's to address these challenges and offer efficient support to decision maker, namely Business Intelligence platforms. Nowadays, the expression "Business Intelligence" is widely spread, and anyone one has a more or less precise idea of what it covers. However, it involves many aspects from most technical ones to very strategic management oriented ones and many authors are tempted to reduce the domain to one or the other perspective. Conversely, industries tend to widen the scope by including analytics, typically, Gartner 2018 report on the domain is entitled "Magic Quadrant for Analytics and Business Intelligence Platforms" [27] . In [13], Chen et al. propose the following definition of Business Intelligence and Analytics (BI&A):

> "[...] BI&A, [...] is often referred to as the techniques, technologies, systems, practices, methodologies, and applications that analyze critical business data to help an enterprise better understand its business and market and make timely business decisions. In addition to the underlying data processing and analytical technologies, BI&A includes business-centric practices and methodologies [...]."

Indeed, a BI platform is conceived to provide access to specific information required by managers in their decision-making process. Consequently, it would be a non-sense to imagine that a BI platform could be developed independently of a deep knowledge of the specific business to which it is dedicated and its strategy. A BI platform as to be seen as one of the technical bricks into the complete wall of the specific business performance management of the concerned company.

To be aligned to the technology focus (the "T") of ICDSST conference, we will limit our purpose in this paper to technical aspect. However, regarding managerial perspective, we would refer the interested readers to the broad literature on Key Performance Indicators (KPI, see for example [52]), scorecards [31–33], and Business Performance Management (BPM, see for example [46, 47]).

From a user perspective, a BI platform appears as framework providing access to a variety of tools among which one commonly find (a selection among): interactive dashboards, reports, OLAP query tools, data mining tools, alerts, ...

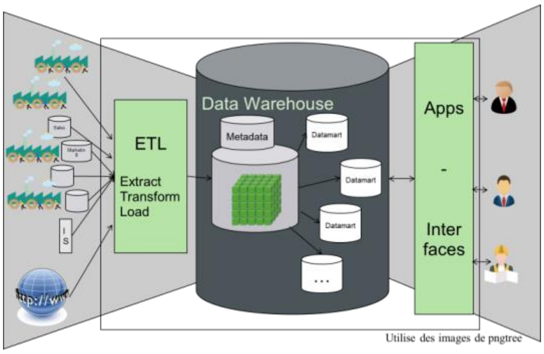
But the specific value of a BI platform lies not so much in the tools offered to users (some of which have existed since long before BI was mentioned) as in the information to which it gives access. Indeed, one of the main challenges of BI is to give users access to information across the organization by allowing them to query a single source while the data needed to build this information is disseminated in multiple sources with heterogeneous structures and riddled with semantic and quality problems. This while avoiding disrupting the behaviour of the operational system.

To achieved this magic, the common reference architecture is structured as illustrated in figure 1. On the right are the users, at the different levels of the organization, who access the relevant applications according to their function. On the left side, the multiple heterogeneous data sources both within and outside the organization are illustrated. In between is the BI platform, the various components of which will be discussed below.

For now, let us underline a few essential aspects expected from such an implementation. As a first specificity, observe that the data flow is one way from operational systems to the analytical system. This is also described as isolation of both world. Combined with adequate data extraction planning, this system

ICDSST 2019 PROCEEDINGS – ONLINE VERSION
THE EWG-DSS 2019 INTERNATIONAL CONFERENCE ON DECISION SUPPORT SYSTEM TECHNOLOGY
P.S. Freitas, F. Dargam, R. Ribeiro, J.M. Moreno, J. Papathanasiou (editors)
Madeira, Portugal, 27-29 May 2019

Fig. 1. Standard Business Intelligence Architecture



splitting implements the requirement that, regardless of the analysis workload, the operational system should not be disrupted.

Downstream of the data flow, all the applications offered to users consult a single source: the data warehouse. Ignoring here the various implementation possibilities (which are discussed below), note that the data warehouse offers a single view of all the retained information, a property generally required under the expression "single truth".

Although they may look like traditional databases, the uses of data warehouses have a set of specificities with respect to operational data bases that impact their design:

- data warehouses do not support update and delete operations, but only inserts (usually batch at night or week-end) and read;
- queries on data warehouses commonly address wide set of data (in lines and columns);
- numbers of users and queries are less limited;
- for huge queries, acceptable response time is higher.

These characteristics gave leads to multidimensional modelling preferably over relational modelling.

Offering a data warehouse requires a sophisticated preliminary integration task denoted Extract – Transform – Load (ETL). These are the three phases of a very complex process of collecting the data, then cleaning, integrating and (re-)formatting before loading in the data warehouse. Depending on the size, availability, variety, and quality of sources, ETL's implementation can be carried out in a wide variety of architectures relying or not on specific data storage.

Finally, it is to note that, even for mature disciplines as data mining or machine learning, being involved interaction with a BI platform offer to the analyst the opportunity to focus their effort on their specific added value, by being largely relieved of data pre-processing issues.

In the rest of the current paper, we aim to offer a double view on BI. First, looking back over the last thirty years, we offer an overview of the standard BI architecture, and the dominant approaches to BI. In a second step, we look to the future and draw an overview of the main challenges BI has to face. Section 2 presents the data warehouse approaches of Inmon and Kimball. Despite their different approaches, they are almost unanimously considered has the fathers of data warehouses, the key component at the heart of standard BI architectures. Then, section 3 extends the technical purpose to the global architecture of BI platforms and presents a typology that organizes their heterogeneity. Turning to the development methodology, section 4 addresses a few BI engineering aspects. After these sections 2 to 4 drawing the state-of-the-art basic BI platform knowledge, section 5 sketches out the main challenges addressed by current scientific literature on BI and section 6 discusses these specifically risen by Big Data.

2 Inmon and Kimball's approaches of data warehouses

Unanimously mentioned as the fathers of data warehouses in both industrial and scientific literature, Inmon and Kimball do not, however, propose completely the same approach.

Inmon defines a data warehouse as "a subject-oriented, integrated, non-volatile, and time-variant collection of data in support of management's decisions" [28]. His approach is commonly qualified as "top-down". He conceives the data warehouse as a single centralised information repository for the entire company at a low level of granularity. The main purpose is to offer a view on data being (i) single truth, (ii) enterprise wide, and (iii) persistent. The implementation of a data warehouse in this perspective is a relational data base. Deduced from the data warehouse a set of departmental data marts are then built to address efficiently the specific needs for reporting and analytics and OLAP queries.

Adopting an approach often consider as more pragmatic, Kimball defines a data warehouse as "a copy of transaction data specifically structured for query and analysis" [34]. His focus relies mainly in the ability to provide efficiently an answer to the actual questions of business management. Kimball build one by one multi-dimensional data marts. Their consistency is guaranteed by "conformed dimensions" which ensure the unicity of a global logical schema as data warehouse.

Beside these two references approaches, a wide variety of implementations and implementations process emerged from the diversity of businesses specificities and projects environments. Extending the scope, the following section presents different BI platform architectures.

3 Global architeture typology

Data warehouses are recognised as key components of BI infrastructure. To be implemented actually and effectively implemented, they need to be integrated into a complete BI infrastructure.

As mentioned above, data warehouses offer a solution to data access, reconciliation and quality problems. To reach that goal, and feed a data warehouse, a significant job has to be done by a commonly called Extract-Transform-Load (ETL) Process.

Downstream of the data warehouse, a whole set of applications with different levels of interactivity are grafted: from predefined reports to interactive dashboards, involving OLAP querying tools or even complete data mining suites.

Each company has its specific IT environment, data sources complexity, heterogeneity, diversity, and update frequencies vary significantly from one business to another. Correlatedly, a variety of architectures can be observed in BI platform architecture implementations. [25] organises a typology following the number of physical data layers.

The single-layer architecture involves no other data storage than the operational data sources. There is no physical data warehouse but a "data warehouse-like" conceptual model which serves as a middleware to access the data sources. This kind of solutions fail to meet the recommended isolation between operational and analytical application and most of the advantages of an actual data warehouse implementation. However, it makes sense to consider such an architecture if the number of data sources is limited, their structures simple and their quality good and if, in addition, the number of analytical queries is relatively low.

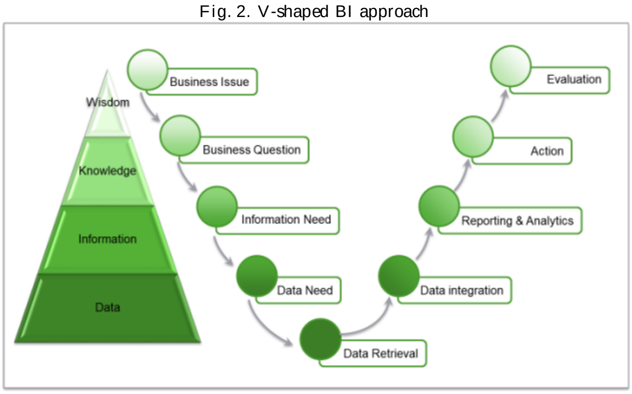
In two layers' architectures, a data warehouse layer is actually implemented. It stores integrated data provided by the ETL process. Its physical implementation involves either an integrated data warehouse, or a set of conformed data marts, or both of them. These architectures are probably the most common in textbooks and offer both the single truth (integration) and the no disruption of operational system qualities commonly expected from BI platforms.

There are situations, in particular when ETL is highly sophisticated, or original data source access very constrained, where ETL process require a specific data storage. A third layer, reconciled data layer is then added. This layer materialized (partially) reconciled source data, not yet fully formatted (integrated, cleaned or whatever required pre-treatment) to be integrated in the data warehouse.

4 BI Engineering

Introduction of BI through the architecture could give the illusion that BI platforms are built in a fully bottom-up approach. This would be a false idea. Indeed, even more critically than for any IT system, BI platforms require a strong strategy/IT alignment [62]. To ensure that the proposed solution meets business needs, a V-shaped approach can be used, as illustrated in Figure 2. In the downward phase, the business question is translated step by step in an information, and a data question. Then, when useful data is retrieved, in the ascending phase of the process, data is gradually aggregated in information and then in knowledge to answer the business question and be integrated as decisions and actions in the managerial process.

Fig. 2. V-shaped BI approach



However, the infrastructure is not developed business question by business question, but a global solution is designed to address wide sets of parametrical requests. This requires strong methodologies. Discussing all the full project management alternatives, would bring us deep in the general project management literature and far out of the allowed length for this paper (interested reader can refer to [44]). Let us focus here on the data warehouse design.

A broad part of industrial literature addresses the data warehouse design at the logical level, starting with a dimensional model as a star or a snowflake schema. This approach can be compared with a data base design that would start at the logical level, designing a relational schema. If this approach can be efficient for small or standard situation, the IT engineering literature has evidenced the need for a preliminary conceptual design drawing entity-relationship schema for example. No language has yet emerged as a standard, the scientific literature involves several propositions to address conceptual data warehouse modelling.

Among them, let us mention [25] and [60]. They both propose similar design processes. The firsts steps, led in parallel, consist, on the one hand, in data sources analysis (bottom-up) and, on the other hand, in requirements analysis (top-down). Then the design goes through conceptual,logical and physical phases. At the conceptual level, Golfarelli and Rizzi [25] propose a language called Dimensional Fact Model (DFM), and Vaisman and Zymanyi [60], a language called MultiDim. Both involves the ability to define facts, dimensions and hierarchies with sophisticated structures.

5 Big Data Challenges

Big Data is one of these buzz words commonly used with a poor and limited understanding. Specifically, beside volume problems, Big Data covers a wide variety of challenges risen by a new generation of data types and sources. Since the emergence Web 2.0, business data has been supplemented by user generated content on social media platforms, integrating semi-structured textual data, pictures and videos. Even more remarkably, e-commerce platforms and multiple apps offer data sources that do not require manual encoding. These sources are constantly multiplying with the omnipresence of mobile devices and will not stop exploding with the deployment of the Internet of Things increasing significantly the volume of spatio-temporal data and frequencies of collection. Let us address them here following the seminal 3Vs' classification: Volume, Velocity, and Variety [21].

5.1 Volume

In the first time of Big Data, some actors have sometimes been tempted to use the term Big Data in an abusive way as a rebranding of traditional data mining and analytical approaches. However, all business areas are now affected by the multiplication of source and Business Intelligence actors can no longer ignore volume challenges. In particular, e-commerce, digital businesses, international companies as well as public administtrations can no longer be satisfied with the support provided by standard technologies.

At the physical level Business Intelligence can benefit from the recent Big Data architectures and tools [14] : Hadoop, GFS, Chukwa, Mapreduce, NoSQL, HBase, Cassandra. [12] proposes a broad overview of the Big Data techniques and technologies addressing the volume challenges at the levels of data capture and storage, data transmission, data curration, data analysis and data visualisation.

Beside the implementation of BI platform on Big Data architecture, specific data warehouse designs aim to target high volume capacity, large data sets while preserving adaptability. One significant attempt in this perspective lies in Data Vault [39] and Conceptual Data Vault [30].

5.2 Velocity

The second "V" of Big Data leads to Velocity. Two challenges are covered by this single word. On the hand, the velocity of porduction of large and complex data, as for exemple streams of images, requires specific real-time pre-processing to be stored in a manageable space and format. On the other hand, systems are expected to provide real-time information. Basic architecture, as discribed above, are not concived to offer hard real-time access to data, solutions can be developped integrating approaches such as those of Business Activity Monitoring (BAM) [26]. For many business, the constraint is (a bit) more flexible. Approaches are the called "almost-real time" or "Just in time" [6,51]. Applications can be found in domains as airlines companies [4], supply chain analytics [57] or on production lines with the industy 4.0.

5.3 Varity

Variety is probably not the least of the challenges involved in Big Data management. Indeed, while businesses are now well doing with standard relational data, Big Data integrates a large set of diversity in data sources formats. Thinking only to the social web, one find unstructured texts, photos, videos, graphs (models for networks), geographic information to mention only the most obvious one. Each of these format brings a set of methods and application domaines: text mining, natural language processing (NLP), sentiment analysis, opinion mining, multilingual analysis, network analysis. Integration of these information in datawarehouses and the analysics tools in BI platform requires new models.

Mobile devices generate large sets of context aware and spatio-temporal data [17]. Specific storage and design methods have been proposed to deal with these in data warehouses [40].

With regard to texts, multidimensional models for OLAP queries on textual documents have been proposed [59], as well as specific tools for semi-structured text, as XML document [53], or on the basis of text metadata or extracted information as keywords [54], topic analysis [63] ,or context [49,50]. Complementarily, specific visualisation and interaction for texts are still an on going challenge [15].

6 Other Challenges

Beside the integegration of Big Data challenges, Business Intelligence scientific litterature addresses many other challenges, this last section before conclusion aims to present two of them that will complete our survey of the fields : the domains addressed by BI and its methodologies.

6.1 Specific domain

Business Intelligence enters all business domains, for many of them this requires specific models development or sytems adaptation as well as the identification and use of levers specific to the business domain.

Non surprisingly, one find among the domain specific studies: market intelligence [58], magazine distribution [41], banking [43], resource allocation [37], production management [38], security and risk management [?, 5], Operational risk management with real-time business intelligence food traceability [23].

Health data and health management are well developped. Specificaly, let us mention Electronic Health Record [8] , radiology [18], health institution management [7, 42], public health management [55], and propositions for a BI implementation framework in healthcare [20].

A last trends to mention proposes a BI system to support decisions about use and evolution of a BI system : BI4BI [9].

6.2 Methodology

The huge litterature on Information Systems design and implementation methodologies provides a solid background for Business Intelligence Platform development. However, specificities of analytical perspective give rise to the development of specific design and implementation strategies [22] as well as specific requirements and goals elicitation methods for the BI platform [11], or one specific component as its data mining tools [10, 48] or the datawarehouse [24]. Users' are also involved to assess the usability of the proposed BI solution [29].

Self-service BI is a recent trends, higly implies users not only clients but elevates them to the rank of co-developers giving (some of) them the ability to develop and integrate their own requests [1–3].

Given the complexity of the systems, the number and diversity of data sources, the sophistication of integration, there is a high risk for BI platform of lack of flexibility and adaptability. An important litterature addresses the challenge of agile development of BI platforms [16, 19, 35, 36, 45, 56].

7 Conclusion

This paper first proposes a standard Business Intelligence approach. It then describes the main technical challenges addressed in the litterature with a particular focus on those risen by the emergence of Big Data.

The Managerial challenges is another aspect of BI which cannot be dissociated of technics in a BI project, [61] explores how business value can be obtained from BI Systems, Trieu summarizes the state of the art in a framework for business value creation from BI that integrates findings. Among the proposals for a complementary research agenda, he identifies: probabilistic models linked necessary condition from BI investments to BI assests to BI impacts; focus on team, industry, and societal levels as well as multi-level studies.

The paper neither addresses security and GDPR aspects which are crooscutting issues across all IT platforms, but which probably require special attention in systems which are developped to provide information in an easy to understand, business oriented fashion.

Employment agencies and job markets never stop underlying the need for competencies in BI, analytics, and Big Data. More and more degrees, in diverse faculties, involve analytical skills. Profiles able to manage the design and develop BI platform can only be developped through a transdisciplinary approach including math, data mining, IT and business perspectives. We hope that this paper could fruitfully contribute to this kind of training by offering to students and researchers in these discipline an introduction to the issues and challenges of Business Intelligence.

References

1. Peggy a Schlesinger and Nayem Rahman. Self-service business intelligence resulting in disruptive technology. *Journal of Computer Information Systems*, 56(1):11–21, 2016.
2. Alberto Abelló, Jérôme Darmont, Lorena Etcheverry, Matteo Golfarelli, Jose-Norberto Mazón, Felix Naumann, Torben Pedersen, Stefano Bach Rizzi, Juan Trujillo, Panos Vassiliadis, and Gottfried Vossen. Fusion cubes: Towards self-service business intelligence. *International Journal of Data Warehousing and Mining (IJDWM)*, 9, 2013.
3. Paul Alpar and Michael Schulz. Self-service business intelligence. *Business & Information Systems Engineering*, 58(2):151–155, Apr 2016.
4. Ron Anderson-Lehman, Hugh J. Watson, Barbara Wixom, and Jeffrey A. Hoffer. Continental airlines flies high with real-time business intelligence. *MIS Quarterly Executive*, 3, 2004.
5. B. Azvine, Z. Cui, B. Majeed, and M. Spott. Operational risk management with real-time business intelligence. *BT Technology Journal*, 25(1):154–167, Jan 2007.
6. B. Azvine, Z. Cui, and D. D. Nauck. Towards real-time business intelligence. *BT Technology Journal*, 23(3):214–225, Jul 2005.
7. Daniele Barone, Thodoros Topaloglou, and John Mylopoulos. Business intelligence modeling in action: A hospital case study. In Jolita Ralyté, Xavier Franch, Sjaak Brinkkemper, and Stanislaw Wrycza, editors, *Advanced Information Systems Engineering*, pages 502–517, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.
8. Wilfred Bonney. Applicability of business intelligence in electronic health record. *Procedia - Social and Behavioral Sciences*, 73:257 – 262, 2013. Proceedings of the 2nd International Conference on Integrated Information (IC-ININFO 2012), Budapest, Hungary, August 30 – September 3, 2012.
9. Manel Brichni, Sophie Dupuy-Chessa, Lilia Gzara, Nadine Mandran, and Corinne Jeannet. Bi4bi: A continuous evaluation system for business intelligence systems. *Expert Systems with Applications*, 76(C):97–112, June 2017.
10. Paola Britos, Oscar Dieste, and Ramón García-Martínez. Requirements elicitation in data mining for business intelligence projects. In David Avison, George M. Kasper, Barbara Pernici, Isabel Ramos, and Dewald Roode, editors, *Advances in Information Systems Research, Education and Practice*, pages 139–150, Boston, MA, 2008. Springer US.
11. Corentin Burnay, Ivan Jureta, Isabelle Linden, and Stéphane Faulkner. A framework for the operationalization of monitoring in business intelligence requirements engineering. *Software and Systems Modeling*, 2014.
12. C.L. Philip Chen and Chun-Yang Zhang. Data-intensive applications, challenges, techniques and technologies: A survey on big data. *Information Sciences*, 275:314 – 347, 2014.
13. Hsinchun Chen, Roger H. L. Chiang, and Veda C. Storey. Business intelligence and analytics: From big data to big impact. *MIS Q.*, 36(4):1165–1188, December 2012.

14. Min Chen, Shiwen Mao, and Yunhao Liu. Big data: A survey. *Mobile Networks and Applications*, 19(2):171–209, Apr 2014.
15. Antoine Clarinval, Isabelle Linden, Anne Wallemacq, and Bruno Dumas. In *Proceedings of the 2018 ACM Symposium on Document Engineering*, United States, 2018. ACM Press.
16. Ken W. Collier. *Agile Analytics: A Value-Driven Approach to Business Intelligence and Data Warehousing*. Addison-Wesley Professional, 1st edition, 2011.
17. Christian Colot, Isabelle Linden, and Philippe Baecke. A survey on mobile data uses. *Intenational Journal of Decision Support System Technologies*, 8(2):29–49, 6 2016.
18. Tessa S. Cook and Paul Nagy. Business intelligence for the radiologist: Making your data work for you. *Journal of the American College of Radiology*, 2014.
19. S. Dasgupta and V. K. Vankayala. Developing real time business intelligence systems the agile way. In *2007 1st Annual IEEE Systems Conference*, pages 1–7, April 2007.
20. Neil Foshay and Craig Kuziemsky. Towards an implementation framework for business intelligence in healthcare. *International Journal of Information Management*, 34(1):20–27, February 2014.
21. Amir Gandomi and Murtaza Haider. Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2):137 – 144, 2015.
22. G. R. Gangadharan and S. N. Swami. Business intelligence systems: design and implementation strategies. In *26th International Conference on Information Technology Interfaces, 2004.*, volume 1, pages 139–144, June 2004.
23. Maria Gianni, Katerina Gotzamani, and Isabelle Linden. How a bi-wise responsible integrated management system may support food traceability. *Int. J. Decis Support Syst. Technol.*, 8(2):1–17, April 2016.
24. Paolo Giorgini, Stefano Rizzi, and Maddalena Garzetti. Goal-oriented requirement analysis for data warehouse design. In *Proceedings of the 8th ACM International Workshop on Data Warehousing and OLAP*, DOLAP '05, pages 47–56, New York, NY, USA, 2005. ACM.
25. Matteo Golfarelli and Stefano Rizzi. *Data Warehouse Design: Modern Principles and Methodologies*. McGraw-Hill, Inc., New York, NY, USA, 1 edition, 2009.
26. Matteo Golfarelli, Stefano Rizzi, Viale Risorgimento, Iuris Cella, and Via Matteotti. Beyond data warehousing: what's next in business intelligence. In *In Proceedings of the 7th ACM international workshop on Data warehousing and OLAP, DOLAP '04*, pages 1–6, 2004.
27. Cindi Howson, Rita Sallam, James Richardson, Joao Tapadinhas, Carlie Idoine, and Alys Woodward. Magic quadrant for analytics and business intelligence platforms. Technical report, 2018.
28. William H. Inmon. *Building the Data Warehouse*. John Wiley & Sons, Inc., New York, NY, USA, 1992.
29. Chrisna Jooste, Judith van Biljon, and Jan Mentz. Usability evaluation for business intelligence applications: A user support perspective. *South African Computer Journal*, 53(0), 2014.
30. Vladan Jovanovic and Ivan Bojicic. Conceptual data vault model. In *Proceedings of the Southern Association for Information Systems Conference, Atlanta, GA, USA March 23rd-24th, 2012*, pages 131–136, 2012.
31. R. S. Kaplan and D. P. Norton. *The Balanced Scorecard: Translating Strategy into Action*. Harvard Business School Press, Boston, 1996.

32. R.S. Kaplan and Norton. *The Strategy-focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment*. The Strategy-focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment. Harvard Business School Press, 2001.
33. R.S. Kaplan and D.P. Norton. *Strategy Maps: Converting Intangible Assets Into Tangible Outcomes*. Harvard Business School Press, 2004.
34. R. Kimball. *The Data Warehouse Toolkit: Practical Techniques for Building Dimensional Data Warehouses*. Wiley, 1996.
35. R. Krawatzeck, B. Dinter, and D. A. P. Thi. How to make business intelligence agile: The agile bi actions catalog. In *2015 48th Hawaii International Conference on System Sciences*, pages 4762–4771, Jan 2015.
36. Deanne Larson and Victor Chang. A review and future direction of agile, business intelligence, analytics and data science. *International Journal of Information Management*, 36(5):700 – 710, 2016.
37. Isabelle Linden. Proposals for the integration of interactive dashboards in business process monitoring to support resources allocation decisions. *Journal of Decision Systems*, 23(3):318–332, 7 2014.
38. Isabelle Linden, Shaofeng Liu, Jonathan Moizer, Pushpa Subramaniam, and Mike Leat. Decision support for erp-based production management: a business intelligence perspective. In *Proceedings of the 15th IFIP WG8.3 International Conference on Decision Support Systems*, 2010.
39. Dan Linstedt and Michael Olschimke. *Building a Scalable Data Warehouse with Data Vault 2.0*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1st edition, 2015.
40. Elzbieta Malinowski and Esteban Zimnyi. *Advanced Data Warehouse Design: From Conventional to Spatial and Temporal Applications (Data-Centric Systems and Applications)*. Springer Publishing Company, Incorporated, 1 edition, 2008.
41. Neil McBride. Business intelligence in magazine distribution. *International Journal of Information Management*, 34(1):58 – 62, 2014.
42. Tobias Mettler and Vivian Vimarlund. Understanding business intelligence in the context of healthcare. *Health Informatics Journal*, 15(3):254–264, 2009. PMID: 19713399.
43. Sérgio Moro, Paulo Cortez, and Paulo Rita. Business intelligence in banking: A literature analysis from 2002 to 2013 using text mining and latent dirichlet allocation. *Expert Systems with Applications*, 42(3):1314 – 1324, 2015.
44. L.T. Moss and Atre S. *Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications*. 2003.
45. Mihaela Muntean. Agile bi – the future of bi. *INFORMATICA ECONOMICA JOURNAL*, volume 3:114–124, 09 2013.
46. A. Neely. *Business Performance Measurement: Unifying Theory and Integrating Practice*. Cambridge University Press, 2007.
47. A. Neely, Inc Books24x7, and R.D. Austin. *Business Performance Measurement: Theory and Practice*. Business Performance Measurement: Theory and Practice. Cambridge University Press, 2002.
48. Celina M. Olszak and Ewa Ziemba. Approach to building and implementing business intelligence systems. *Interdisciplinary Journal of Information, Knowledge, and Management*, 2:135–148, 2007.
49. Lamia Oukid, Nadjia Benblidia, Fadila Bentayeb, Ounas Asfari, and Omar Boussaid. Contextualized text olap based on information retrieval. *International Journal of Data Warehousing and Mining*, 11:1–21, 04 2015.

50. Lamia Oukid, Omar Boussaid, Nadjia Benblidia, and Fadila Bentayeb. TLabel: A New OLAP Aggregation Operator in Text Cubes. *International Journal of Data Warehousing and Mining*, 12(4):54–74, 2016.
51. Zeljko Panian. Just-in-time business intelligence and real-time decisioning. In *AIC'09 Proceedings of the 9th WSEAS international conference on Applied informatics and communications*, pages 106–111.
52. D. Parmenter. *Key Performance Indicators: Developing, Implementing, and Using Winning KPIs*. Wiley, 3rd edition, 2015.
53. Geneviève Pujolle, Franck Ravat, Olivier Teste, Ronan Tournier, and Gilles Zurfluh. Multidimensional database design from document-centric xml documents. In Alfredo Cuzzocrea and Umeshwar Dayal, editors, *Data Warehousing and Knowledge Discovery*, pages 51–65, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.
54. Franck Ravat, Olivier Teste, Ronan Tournier, and Gilles Zurfluh. Top_keyword: An aggregation function for textual document olap. In Il-Yeol Song, Johann Eder, and Tho Manh Nguyen, editors, *Data Warehousing and Knowledge Discovery*, pages 55–64, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.
55. Sonia Rivest, Yvan Bédard, Marie-Josée Proulx, Martin Nadeau, Frederic Hubert, and Julien Pastor. Solap technology: Merging business intelligence with geospatial technology for interactive spatio-temporal exploration and analysis of data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60(1):17 – 33, 2005. Including Theme Section:.
56. Mohamad Sadegh Sangari and Jafar Razmi. Business intelligence competence, agile capabilities, and agile performance in supply chain an empirical study. *The International Journal of Logistics Management*, 26, 08 2015.
57. B.S. Sahay and Jayanthi Ranjan. Real time business intelligence in supply chain analytics. *Information Management & Computer Security*, 16(1):28–48, 2008.
58. Merlin David Stone and Neil David Woodcock. Interactive, direct and digital marketing: A future that depends on better use of business intelligence. *Journal of Research in Interactive Marketing*, 8(1):4–17, 2014.
59. Frank S. C. Tseng and Annie Y. H. Chou. The concept of document warehousing for multi-dimensional modeling of textual-based business intelligence. *Decision Support System*, 42(2):727–744, November 2006.
60. Alejandro Vaisman and Esteban Zimányi. *Data Warehouse Systems: Design and Implementation*. Springer, Heidelberg, 2014.
61. Trieu Van-Hau. Getting value from business intelligence systems: A review and research agenda. *Decision Support Systems*, 93:111 – 124, 2017.
62. J.M. Villamarín García and B.H. Díaz Pinzón. Key Success Factors to Business Intelligence Solution Implementation. *Journal of Intelligence Studies in Business*, 7(1):46–69, 2017.
63. Duo Zhang, ChengXiang Zhai, Jiawei Han, Ashok Srivastava, and Nikunj Oza. Topic modeling for olap on multidimensional text databases: topic cube and its applications. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 2(5-6):378–395.

EmC-ICDSST 2019
5th International Conference on Decision Support System Technology – ICDSST 2019 & EURO Mini Conference 2019

27 – 29 May 2019 – University of Madeira, Funchal, Madeira, Portugal

ISBN: 978-989-8805-4-7