MODELLING AND SIMULATION 2019

THE EUROPEAN SIMULATION

AND

MODELLING CONFERENCE

2019

ESM®'2019

EDITED BY

Pilar Fuster-Parra

and

Óscar Valero Sierra

OCTOBER 28-30, 2019

PALMA DE MALLORCA

SPAIN

A Publication of EUROSIS-ETI

The 33rd Annual European Simulation and Modelling Conference 2019

PALMA DE MALLORCA, SPAIN

OCTOBER 28-30, 2019

Organised by

ETI - The European Technology Institute

Sponsored by

EUROSIS - The European Simulation Society

Ghent University

Co-Sponsored by

University of Skovde

University of Žilina

GODAN

Hosted by

Universitat de les Illes Balears
Palma de Mallorca, Spain

EXECUTIVE EDITOR

PHILIPPE GERIL
(BELGIUM)

EDITORS

General Conference Chairs

Pilar Fuster-Parra
Universitat de les Illes Balears
Palma de Mallorca, Spain
Óscar Valero
Universitat de les Illes Balears
Palma de Mallorca, Spain

Local Programme Committee

José Guerrero, Universitat de les Illes Balears, Palma de Mallorca, Spain
Javier Martín, Universitat de les Illes Balears, Palma de Mallorca, Spain
Juan José Miñana, Universitat de les Illes Balears, Palma de Mallorca, Spain
Carles Mulet, Universitat de les Illes Balears, Palma de Mallorca, Spain
Óscar Valero, Universitat de les Illes Balears, Palma de Mallorca, Spain
Pilar Fuster-Parra, Universitat de les Illes Balears, Palma de Mallorca, Spain

Past Conference Chair
Dieter Claeys
Ghent University
Zwijnaarde (Gent) and Flanders Make, Belgium
Veronique Limère
Ghent University
Ghent, Belgium

Journal Publication Chairs
Yan Luo, NIST, Gaithersburg, USA
Peter Lawrence, Swinburne University, Australia
Dr. Wan, International Islamic University Malaysia

ESM Conference Chair
António Carvalho Brito, FEUP - University of Porto, Porto, Portugal

INTERNATIONAL PROGRAMME COMMITTEE

Methodology and Tools
Claudia Krull, Otto-von-Guericke University, Magdeburg, Germany
J. Manuel Feliz Teixeira, University of Porto, Porto, Portugal
Jerzy Respondek, Silesian University of Technology, Gliwice, Poland

Discrete Simulation Modeling Techniques and Tools
Helge Hagenauer, Universitaet Salzburg, Salzburg, Austria
Sophie Hennequin, ENIM, Metz Cedex, France

Simulation and Artificial Intelligence
Helder Coelho, Fac Ciencias, Lisbon, Portugal
Paulo Cortez, University of Minho, Guimareas, Portugal
Martin Hruby, Brno University of Technology, Brno, Czech Republic
Vladimir Janousek, Brno University of Technology, Brno, Czech Republic
Leon Rothkrantz, Prague University of Technology, Prague, Czech Republic

INTERNATIONAL PROGRAMME COMMITTEE

AI and Fuzzy Systems
Tomasa Calvo, Universidad de Alcalá de Henares, Madrid, Spain
Pilar Fuster-Parra, Universitat de les Illes Balears, Palma de Mallorca, Spain
José Guerrero, Universitat de les Illes Balears, Palma de Mallorca, Spain
Javier Martín, Universitat de les Illes Balears, Palma de Mallorca, Spain
Juan José Miñana, Universitat de les Illes Balears, Palma de Mallorca, Spain
Carles Mulet, Universitat de les Illes Balears, Palma de Mallorca, Spain
Ali Shams Nateri, University of Guilan, Rasht, Iran
Oscar Valero, Universitat de les Illes Balears, Palma de Mallorca, Spain

Agent Based Simulation
Kurt De Cock, Ghent University, Ghent, Belgium
Pilar Fuster-Parra, Universitat de les Illes Balears, Palma de Mallorca, Spain
José Guerrero, Universitat de les Illes Balears, Palma de Mallorca, Spain
Ioan Alfred Letia, TU Cluj Napoca, Romania
Javier Martín, Universitat de les Illes Balears, Palma de Mallorca, Spain
Juan José Miñana, Universitat de les Illes Balears, Palma de Mallorca, Spain
Carles Mulet, Universitat de les Illes Balears, Palma de Mallorca, Spain
Isabel Praca, Ist. Superior do Porto, Portugal
Oscar Valero, Universitat de les Illes Balears, Palma de Mallorca, Spain

Simulation and Optimization
José António Oliveira, Universidade do Minho, Campus de Gualtar, Braga, Portugal
Janos-Sebestyen Janosy, Hungarian Academy of Sciences, Budapest, Hungary

oT and Smart Industry (Internet of Things & Industry 4.0)
Track Chair
Abderrazak Jemai, INSAT, University of Carthage, Tunis, Tunisia

Ahmed Chiheb Ammari, INSAT, Tunis, Tunisia
Mhamed Ghazel, CNRS, Lille, France
Habib Smei, Iset Rades, Tunis, Tunisia

High Performance Large Scale and Hybrid Computing
Pierre Siron, ONERA, Toulouse, France
Jingjing Wang, SUNY Binghamton University, New York, USA

Simulation in Education and Graphics Visualization
Marco Roccetti, University of Bologna, Italy
Ranjit Singh, UB Patient Safety Research Center, University of Buffalo, Buffalo NY USA

Simulation in Environment, Ecology, Biology and Medicine
Joel Colloc, Université du Havre, Le Havre, France
Laurent Perochon, VetaGro-Sup, Lempdes, France

Analytical and Numerical Modelling Techniques
Ana M. Camacho, UNED, Madrid, Spain
Clemens Heitzinger, TU Vienna, Vienna, Austria

Web and Cloud Based Simulation
Manuel Alfonseca, Universidad Autonoma de Madrid, Spain
Ammar Al-Khani, Nordphysics, Espoo, Finland
Peter Kvasnica, Alexander Dubcek University of Trencin, Trencin, Slovak Republic
Yan Luo, NIST, Gaithersburg, USA
Jose Machado, University of Minho, Braga, Portugal

Physics Modelling and Cosmological Simulation
José Manuel Feliz-Teixeira, University of Porto, Porto, Portugal
Philippe Geril, ETI Bvba, Ostend, Belgium

INTERNATIONAL PROGRAMME COMMITTEE

Information Fusion, Fuzzy Theory and applications to Engineering, Medicine and Economics
Manuel Ojeda Aciego, Universidad de Málaga, Malaga, Spain
Maria Jesús Campion, Universidad de Navarra, Pamplona, Spain
Javier Gutiérrez García, Universidad País Vasco (UPV), Bilbao, Spain
Samuel Morillas Gómez, Universitat Politècnica de València, Valencia, Spain
Esteban Indurain, Universidad de Navarra, Pamplona, Spain
Jesús Rodriguez López, Universitat de València, Valencia, Spain
José M. Merigó, Universidad de Chile, Santiago, Chile
Radko Mesiar, Slovak University of Technology, Bratislava, Slovakia
Javier Montero, Universidad Complutense Madrid, Madrid, Spain
Pilar Fuster-Parra, Universitat de les Illes Balears, Palma de Mallorca, Spain
Jordi Recasens, Universitat Politècnica de Calalunya, Barcelona, Spain
Susana Montes Rodriguez, Universidad de Oviedo, Asturias, Spain
Oscar Valero, Universitat de les Illes Balears, Palma de Mallorca, Spain

Simulation in Energy and Power Systems
Sérgio Leitão, CETAV/UTAD, Vila Real, Portugal
Sergeij Kalashnkikow, DANFOSS, Austria
Janos-Sebestyen Janosy, KFKI Atomic Energy Research Institute, Budapest, Hungary

Simulation in Engineering Processes
Chrissanti Angeli,Technological Institute of Piraeus, Athens, Greece
Bernhard Heinzl, TU Wien, Vienna, Austria
Ali Shams Nateri, University of Guilan, Rasht, Iran
Jan Studzinski, Polish Academy of Sciences, Warsaw, Poland

Simulation in Model Driven Engineering
Paulo Jorge Sequeira Gonçalves, Polytechnic Institute of Castelo Branco, Castelo Branco, Portugal

Intelligent Systems
Ying He, De Montfort University, Leicester, United Kingdom
José Machado, Universidade do Minho, Braga, Portugal
Manuel Filipe Santos, Universidade do Minho, Guimarães, Portugal

Simulation in Hospital Logistics
Joel Colloc, Le Havre Normandy University, Le Havre, France
Jose Machado, University of Minho, Braga, Portugal
Jose Antonio V. Oliveira, University of Minho, Braga, Portugal

Simulation in Logistics
El-Houssaine Aghezzaf, Ghent University, Ghent, Belgium
Remy Dupas, Université de Bordeaux , Bordeaux, France
Olivier Grunder, UTBM, Belfort, France
Marie-Ange Manier, UTBM, Belfort, France
Roberto Razzoli, University of Genova, Genova, Italy
Rosaldo Rossetti, University of Porto, Porto, Portugal
Pengjun Zheng, Ningbo University, Zhejiang, China P.R.

Supply Chain Simulation
Florina Covaci,"Babes-Bolyai" University, Cluj-Napoca, Romania
Eleni Mangina, University College Dublin (UCD), Dublin, Ireland
Sang Do No, Sungkyunkwan University, South Korea

Simulation with Petri Nets
Francois Siewe, De Montfort University, Leicester, United Kingdom
Pascal Berruet, Universite Bretagne Sud, Lorient, France
Stefano Marrone, Seconda Universita degli Studi di Napoli, Naples, Italy
Alexandre Nketsa, LAAS-CNRS, Toulouse, France

INTERNATIONAL PROGRAMME COMMITTEE

Bond Graphs Simulation
Jesus Felez, Univ. Politecnica de Madrid, Spain
Andre Tavernier, BioSim, Brussels, Belgium

DEVS
Fernando Tricas, Universidad de Zaragoza, Spain

Fluid Flow Simulation
H.A.Nour Eldin, University of Wuppertal, Germany
Markus Fiedler, Blekinge Institute of Technology, Sweden

Emergency Risk Management Simulation
Konstantinos Kirytopoulos, University of South Australia, Australia
Lode Vermeersch, Credendo, Brussels, Belgium

All articles published in these Proceedings have been peer reviewed.

EUROSIS-ETI Publications are ISI-Thomson and IET referenced.
ESM® Proceedings are indexed on ISI-Thomson, IET, SCOPUS and Elsevier Engineering Village.

A CIP Catalogue record for this book is available from the Royal Library of Belgium under nr.12620

For permission to publish a complete paper write EUROSIS, c/o Philippe Geril, ETI Executive Director, Greenbridge Science Park, Ghent University – Ostend Campus, Wetenschapspark 1, Plassendale 1, B-8400 Ostend, Belgium.

EUROSIS is a Division of ETI Bvba, The European Technology Institute, Torhoutsesteenweg 162, Box 4.02, B-8400 Ostend, Belgium

Printed in Belgium by Reproduct NV, Ghent, Belgium
Cover Design by C-Design Bvba, Wielsbeke, Belgium

EUROSIS-ETI Publication

ISBN: 978-9492859-09-9
EAN: 9789492859099

EUROPEAN SIMULATION AND MODELLING CONFERENCE 2019

Preface

Dear participants,

It is our honour to welcome each and every one of you to the 33rd European Simulation and Modelling Conference - ESM'2019 held in association with Balearic Islands University. During two and a half days, this conference aims to be a meeting place where scientists and technicians in the field of Simulation and Modelling can find a discussion forum to promote research and exchange ideas. The programme is organized to allow for ample opportunity to meet and discuss, encouraging a high level of interaction and reflection on new research ideas.

We have some 50 participants from 20 countries with presentations grouped in 13 main themes such as Analytical and Numerical Modelling, Simulation Optimization, Parallel Simulation, Simulation using Fuzzy Algorithms, AI Based Data Analysis, Cloud Based Simulation, Internet of Things, Financial Data Simulation, Smart Grid Simulation, Environmental and Population Dynamics Simulation, Medical Simulation, Engineering Simulation, Traffic and Logistics Simulation.

Furthermore, we are honoured to announce our keynote speakers, José M. Merigó (University of Chile and University of Technology of Sydney), talking about the Ordered Weighted Average: theory and applications, José Javier Ramasco (IFISC, Balearic Islands University) talking about Mobile Phone Records to feed Activity-based Travel Demand Models, and José Guerrero (Balearic Islands University) talking about a New Framework for Multi-Agent Systems: A Possibilistic Viewpoint. A tutorial by Janos Sebestyen Janosy (Centre for Energy Research, Hungarian Academy of Sciences) about Consolidating unpredictable power sources – SMARTGRID would be presented. We are very grateful that they could accept our invitation to share their expertise with us.

We are indebted to a number of people without whom the conference would not be possible. First of all, we would like to thank all authors for sharing their contributions. Our thanks also go to all members of the International Program Committee for the critical reviewing work that was key to maintaining the high scientific quality of the conference.

Furthermore, we would in particular like to thank Philippe Geril from EUROSIS, whose continued dedication and hard work, as the conference organizer, was indispensable for a successful organization of ESM'2019.

Finally, we would like to wish you all a stimulating and inspiring experience at the conference and a pleasant stay in the beautiful city of Palma de Mallorca.

Pilar Fuster Parra and Óscar Valero Sierra
Dpt. Mathematics and Computer Sciences
Balearic Islands University, Spain

ESM'2019 General Chairs

CONTENTS

CONTENTS

FINANCIAL DATA SIMULATION

SMART GRID SIMULATION

CONTENTS

ENVIRONMENTAL AND POPULATION DYNAMICS SIMULATION

MEDICAL SIMULATION

CONTENTS

ENGINEERING SIMULATION

LOGISTICS SIMULATION

SCIENTIFIC PROGRAMME

ANALYTICAL AND NUMERICAL MODELLING

APPLYING HOMOTOPY PERTURBATION METHOD TO NON-LINEAR DIFFERENTIAL EQUATIONS

Linas Petkevičius
Institute of Computer Science
Vilnius University
Didlaukio str. 47, LT-08303 Vilnius, Lithuania
Email: linas.petkevicius@mif.vu.lt

KEYWORDS
Homotopy perturbation method, non-linear differential equation, reaction-diffusion, reactors, modeling

ABSTRACT

This research investigates the analytical approximations using the homotopy perturbation method. The six different initial linearalizations are considered and it is demonstrated that errors of those approximations might vary by multiple magnitudes of order. It is concluded that different linearization methods for different parameters provide the best analytical approximations. We have also demonstrated that due to the complexity of approximations the first order approximations formulas are actually known analytical solutions under limiting cases.

INTRODUCTION

The Homotopy perturbation method (HPM) was proposed by Ji Huan He (He 1999). The method allow to obtain approximate analytical solutions for non-linear differential equations. HPM is a combination of homotopy in topology and perturbation techniques. The method allows to obtain analytical approximation for a wide variety of models arising in different fields (Filobello-Nino et al. 2019, Fedorov et al. 2019). Some approximations have been provided for the enzyme kinetics models derived from the classical Michaelis-Menten kinetic model (Saranya et al. 2018, Filobello-Nino et al. 2019). However, approximations are suitable on very limited cases duo to the initial linearization of perturbation models which led to different accuracy of analytical approximations. In this paper we investigate six different initial linearizations of perturbation.

Due to the common practice to provide zero or first order approximations of HMP having large expressions, in this work we emphasize attention that those expressions actually are known analytical solutions.

The publication consists of four sections: the mathematical model of bioreactors kinetics is provided, then HPM is used to get analytical approximations, the different linearization models are provided in the following section and results conclude the paper.

MODEL OF BIOREACTORS

Let us consider the bioreactors mathematical model at steady-state (Baronas et al. 2018), the concentration of substrate is $\widetilde{S} = \widetilde{S}(\widetilde{\rho})$, where $\widetilde{\rho}$ is the distance from reactor center. Let us define the inner layer where reaction and diffusion take place by $0 < \widetilde{\rho} < 1$, and external layer where only diffusion takes place by $1 < \widetilde{\rho} < 1 + \widetilde{\nu}$. The governing equations are as follows:

$$D^*(\widetilde{\rho})\Delta\widetilde{S}(\widetilde{\rho}) = a^*(\widetilde{\rho})\frac{\sigma^2\widetilde{S}(\widetilde{\rho})}{1+\widetilde{S}(\widetilde{\rho})}, \quad (1)$$

$$D^*(\widetilde{\rho}) = \begin{cases} 1, & 0 < \rho < 1, \\ \gamma = \frac{D_{S,n}}{D_{S,m}}, & 1 < \widetilde{\rho} < 1 + \widetilde{\nu}, \end{cases} \quad (2)$$

$$a^*(\widetilde{\rho}) = \begin{cases} 1, & 0 < \widetilde{\rho} < 1, \\ 0, & 1 < \widetilde{\rho} < 1 + \widetilde{\nu}, \end{cases} \quad (3)$$

$$\sigma^2 = \frac{V_{\max}\rho_m^2}{K_M D_{S,m}}, \quad \beta = \frac{\gamma\widetilde{\rho}_1}{\widetilde{\rho}_1 - \widetilde{\rho}_0} = \frac{D_n\widetilde{\rho}_1}{D_m(\widetilde{\rho}_1 - \widetilde{\rho}_0)}. \quad (4)$$

where $\widetilde{\nu}$ denotes dimensionless Nernst layer thickness, D is diffusion coefficient, σ is Thiele module, also known as Damkohler number (Baronas et al. 2018). The dimensionless factor β can be considered as the effective Biot number or dimensionless mass transfer coefficient that quantifies the relative preponderance of internal or external diffusion.

The boundary conditions are defined as follows:

$$\frac{\partial\widetilde{S}}{\partial\widetilde{\rho}}\Big|_{\widetilde{\rho}=1^-} = \frac{D_{S,n}}{D_{S,m}}\frac{\partial\widetilde{S}}{\partial\widetilde{\rho}}\Big|_{\widetilde{\rho}=1^+}, \widetilde{S}(1^-) = \phi\widetilde{S}(1^+), \quad (5)$$

$$\frac{\partial\widetilde{S}}{\partial\widetilde{\rho}}\Big|_{\widetilde{\rho}=0} = 0, \quad \widetilde{S}(1+\widetilde{\nu}) = \widetilde{S}_0. \quad (6)$$

The mathematical model (1)-(6) was consider in (Eswari and Rajendran 2010, Baronas et al. 2018)

APPROXIMATIONS OF SOLUTION

Let us consider different initial equation linearizations as the models (M1)-(M6):

$$\Big(\widetilde{S}''(\widetilde{\rho})+\frac{2}{\widetilde{\rho}}\widetilde{S}'(\widetilde{\rho})\Big)-\varepsilon\frac{\sigma^2\widetilde{S}(\widetilde{\rho})}{1+\widetilde{S}(\widetilde{\rho})}= \quad 0 \quad \text{(M1)}$$

$$\Big(\widetilde{S}''(\widetilde{\rho})+\frac{2}{\widetilde{\rho}}\widetilde{S}'(\widetilde{\rho})\Big)\Big(1+\widetilde{S}(\widetilde{\rho})\Big)-\varepsilon\sigma^2\widetilde{S}(\widetilde{\rho})= \quad 0 \quad \text{(M2)}$$

$$\widetilde{S}''(\widetilde{\rho})+\frac{2}{\widetilde{\rho}}\widetilde{S}'(\widetilde{\rho})-\sigma^2\widetilde{S}(\widetilde{\rho})+\widetilde{S}^{\varepsilon}(\widetilde{\rho})\Big(\widetilde{S}''(\widetilde{\rho})+\frac{2}{\widetilde{\rho}}\widetilde{S}'(\widetilde{\rho})\Big)= \quad 0 \quad \text{(M3)}$$

$$\Big(\widetilde{S}''(\widetilde{\rho})+\frac{2}{\widetilde{\rho}}\widetilde{S}'(\widetilde{\rho})\Big)-\sigma^2+\varepsilon\frac{\sigma^2}{1+\widetilde{S}(\widetilde{\rho})}= \quad 0 \quad \text{(M4)}$$

$$\Big(\widetilde{S}''(\widetilde{\rho})+\frac{2}{\widetilde{\rho}}\widetilde{S}'(\widetilde{\rho})\Big)\Big(1+\varepsilon\widetilde{S}(\widetilde{\rho})\Big)-\sigma^2\widetilde{S}(\widetilde{\rho})= \quad 0 \quad \text{(M5)}$$

A small parameter ϵ in equations (M2) and (M5) can be found in literature (Eswari and Rajendran 2010), the models (M1), (M3), (M4) are not considered. Put $S=\sum_{i=0}^{\infty}u_i(r)\epsilon^i$ to (M1)-(M5) and collecting members near power of ε , then solve the differential equations (1)-(6) (Bender and Orszag 2013).
The members near powers of member (M1):

$$\begin{aligned}
\varepsilon^0 \quad &: \quad \Big(\widetilde{S}_0''(\widetilde{\rho})+\frac{2}{\widetilde{\rho}}\widetilde{S}_0'(\widetilde{\rho})\Big)=0,\\
\widetilde{S}_{M1,0}(\widetilde{\rho}) \quad &= \quad \begin{cases}\phi S_0, & 0<\widetilde{\rho}<1,\\ S_0, & 1<\widetilde{\rho}<1+\widetilde{\nu},\end{cases}\\
\varepsilon^1 \quad &: \quad \Big(\widetilde{S}_1''(\widetilde{\rho})+\frac{2}{\widetilde{\rho}}\widetilde{S}_1'(\widetilde{\rho})\Big)-\frac{\sigma^2\widetilde{S}_0(\widetilde{\rho})}{1+\widetilde{S}_0(\widetilde{\rho})}=0,\\
\widetilde{S}_{M1,1}(\widetilde{\rho}) \quad &= \quad \begin{cases}\frac{1}{6}\frac{\sigma\phi S_0}{\gamma(1+\widetilde{\nu})(\phi S_0+1)}((1+\widetilde{\nu})\gamma\widetilde{\rho}^2-2\rho_N\phi-\rho_N\gamma+2\phi),\\ \frac{1}{3}\frac{\sigma\phi S_0}{\gamma(1+\widetilde{\nu})(\phi S_0+1)}(1-\frac{(1+\widetilde{\nu})}{\widetilde{\rho}}),\end{cases}\\
\varepsilon^2 \quad &: \quad \Big(\widetilde{S}_2''(\widetilde{\rho})+\frac{2}{\widetilde{\rho}}\widetilde{S}_2'(\widetilde{\rho})\Big)-\frac{\sigma^2\widetilde{S}_1(\widetilde{\rho})}{1+\widetilde{S}_0(\widetilde{\rho})}\\
&\quad +\frac{\sigma^2\widetilde{S}_0(\widetilde{\rho})\widetilde{S}_1(\widetilde{\rho})}{(1+\widetilde{S}_0(\widetilde{\rho}))^2}=0\\
\widetilde{S}_{M1,2}(\widetilde{\rho}) \quad &= \quad \begin{cases}\frac{1}{8\cdot45}\frac{1}{\gamma^2(1+\widetilde{\nu})^2(\phi^3S_0^3+3\phi^2S_0^2+3\phi S_0+1)}\big(\phi S_0\sigma^2(3\gamma^2(1+\widetilde{\nu})^2\widetilde{\rho}^4,\\ -20\gamma(1+\widetilde{\nu})^2\phi\widetilde{\rho}^2+40(1+\widetilde{\nu})^2\phi^2-10\gamma^2(1+\widetilde{\nu})^2\widetilde{\rho}^2\\ +20\gamma(1+\widetilde{\nu})\phi\widetilde{\rho}^2+28\gamma(1+\widetilde{\nu})^2\phi\\ -80(1+\widetilde{\nu})\phi^2+7\gamma^2(1+\widetilde{\nu})^2-28\gamma(1+\widetilde{\nu})\phi+40\phi^2)\big)\\ \frac{1}{45}\frac{\phi S_0\sigma^2(5(1+\widetilde{\nu})\phi+(1+\widetilde{\nu})\gamma-5\phi)}{\gamma^2(1+\widetilde{\nu})^2(\phi^3S_0^3+3\phi^2S_0^2+3\phi S_0+1)}(\frac{(1+\widetilde{\nu})}{\widetilde{\rho}}-1),\end{cases}\\
\varepsilon^3 \quad &: \quad \Big(\widetilde{S}_3''(\widetilde{\rho})+\frac{2}{\widetilde{\rho}}\widetilde{S}_3'(\widetilde{\rho})\Big)-\frac{\sigma^2\widetilde{S}_2(\widetilde{\rho})}{1+\widetilde{S}_0(\widetilde{\rho})}\\
&\quad +\frac{\sigma^2\widetilde{S}_0(\widetilde{\rho})\widetilde{S}_2(\widetilde{\rho})}{(1+\widetilde{S}_0(\widetilde{\rho}))^2}+\frac{\sigma^2\widetilde{S}_1^2(\widetilde{\rho})}{(1+\widetilde{S}_0(\widetilde{\rho}))^2}\\
&\quad -\frac{\sigma^2\widetilde{S}_0(\widetilde{\rho})\widetilde{S}_1^2(\widetilde{\rho})}{(1+\widetilde{S}_0(\widetilde{\rho}))^3}=0\\
&..
\end{aligned}$$

$$\begin{aligned}
\varepsilon^n \quad &: \quad \Big(\widetilde{S}_n''(\widetilde{\rho})+\frac{2}{\widetilde{\rho}}\widetilde{S}_n'(\widetilde{\rho})\Big)-\frac{\sigma^2\widetilde{S}_n(\widetilde{\rho})}{1+\widetilde{S}_0(\widetilde{\rho})}\\
&\quad + \quad \frac{\sigma^2\sum_{i,j\in I_1}\widetilde{S}_j(\widetilde{\rho})\widetilde{S}_i(\widetilde{\rho})}{(1+\widetilde{S}_0(\widetilde{\rho}))^2}+\\
&\quad +\sigma^2\sum_{i=3}^{n-1}(-1)^n\frac{\sum_{i,j\in I_2}\widetilde{S}_j(\widetilde{\rho})\widetilde{S}_i(\widetilde{\rho})}{(1+\widetilde{S}_0(\widetilde{\rho}))^i}\\
&\quad +\frac{(-1)^n\sigma^2\widetilde{S}_0(\widetilde{\rho})\widetilde{S}_1^{n-1}(\widetilde{\rho})}{(1+\widetilde{S}_0(\widetilde{\rho}))^n}=0\\
&\quad \widetilde{S}_{M1,n}(\widetilde{\rho})=\begin{cases}\sum_{i=0}^{n}C_i\widetilde{\rho}^{2i}, & 0<\widetilde{\rho}<1,\\ C_{n+1}+C_{n+2}\frac{1}{\widetilde{\rho}}, & 1<\widetilde{\rho}<1+\widetilde{\nu},\end{cases}
\end{aligned}$$

In an analogous way (M2) - (M5) becomes:

$$\begin{aligned}
\widetilde{S}_{M2,0}(\widetilde{\rho}) \quad &= \quad \begin{cases}\phi S_0\\ S_0\end{cases}\\
\widetilde{S}_{M2,1}(\widetilde{\rho}) \quad &= \quad \begin{cases}\sigma^2\phi S_0\big((1+\widetilde{\nu})\gamma\widetilde{\rho}^2-2(1+\widetilde{\nu})\phi-(1+\widetilde{\nu})\gamma+2\phi\big)/\\ \big(6\gamma(1+\widetilde{\nu})(\phi S_0+1)\big)\\ -\sigma^2\phi S_0((1+\widetilde{\nu})-\widetilde{\rho})/(3\gamma(1+\widetilde{\nu})(\phi S_0+1)\widetilde{\rho})\end{cases}\\
\widetilde{S}_{M3,0}(\widetilde{\rho}) \quad &= \quad \begin{cases}\frac{C_1\sinh((1/2)\sqrt{2}\widetilde{\rho})}{\widetilde{\rho}}\\ C_2+\frac{C_3}{\widetilde{\rho}}\end{cases}
\end{aligned}$$

$$\widetilde{S}_{M4,0}(\widetilde{\rho}) \quad = \quad \begin{cases}\frac{((1+\widetilde{\nu})\gamma\widetilde{\rho}^2\sigma^2+6(1+\widetilde{\nu})\gamma S_0+2\sigma^2)}{(6\gamma(1+\widetilde{\nu}))}\\ C_5+\frac{C_6}{\widetilde{\rho}}\end{cases} \quad (7)$$

$$\widetilde{S}_{M5,0}(\widetilde{\rho}) \quad = \quad \begin{cases}\frac{C_4\sinh(\sigma^2\widetilde{\rho})}{\widetilde{\rho}}\\ C_7+\frac{C_8}{\widetilde{\rho}}\end{cases} \quad (8)$$

where $C_i, i=1,\ldots,8$ are constants that were found by applying boundary conditions. The constants are not provided in publication due to a large expressions.
By looking to (M1) n-th part of solution, let us look for solution in domain $0<\widetilde{\rho}<1$ of the following type:

$$\widetilde{S}_{M6}^{m}(\widetilde{\rho})=\sum_{n=0}^{m}a_n\widetilde{\rho}^{2n} \quad \text{(M6)}$$

By putting (M6) to (1) receive that a_n we have expression of:

$$a_n=\frac{1}{n!}\frac{a_0\varepsilon^n\sum_{i=0}^{n}b_ia_0^i}{\sum_{j=0}^{2n+1}\binom{2n+1}{j}a_0^j} \quad (9)$$

By solving differential equation (1)-(6), the parameters become:

$$\begin{aligned}
a_0 \quad &= \quad \phi\widetilde{S}_0-\frac{1}{6}\sigma^2\phi\widetilde{S}_0\frac{2(1+\widetilde{\nu})\phi+(1+\widetilde{\nu})\gamma-2\phi}{(1+\widetilde{\nu})\gamma(\phi\widetilde{S}_0+1)},\\
a_1 \quad &= \quad \frac{1}{6}\cdot\frac{\varepsilon a_0}{1+a_0}, \quad a_2=\frac{1}{120}\cdot\frac{\varepsilon^2a_0}{a_0^3+3a_0^2+3a_0+1}\\
&..
\end{aligned}$$

Table 1: The approximation error of models (M1) - (M6) at: $S_0 = 1, \phi = 0.5, \sigma^2 = \gamma = \widehat{\nu} = 1$ (left), $S_0 = 100, \phi = 0.5, \sigma^2 = \gamma = \widehat{\nu} = 1$ (right)

Model	0-order	1-order	2-order	0-order	1-order	2-order
(M1)	$1.1 \cdot 10^{-1}$	$3.5 \cdot 10^{-4}$	$1.19 \cdot 10^{-7}$	$9.6 \cdot 10^{-1}$	$2.12 \cdot 10^{-9}$	$5 \cdot 10^{-10}$
(M2)	$1.1 \cdot 10^{-1}$	$3.5 \cdot 10^{-4}$	$1.19 \cdot 10^{-7}$	$9.6 \cdot 10^{-1}$	$2.12 \cdot 10^{-9}$	$\mathbf{1.6 \cdot 10^{-14}}$
(M3)	$\mathbf{7.3 \cdot 10^{-3}}$	$\mathbf{4.3 \cdot 10^{-5}}$	-	$4.4 \cdot 10^{2}$	$1.59 \cdot 10^{2}$	-
(M4)	$1.7 \cdot 10^{-1}$	$7 \cdot 10^{-1}$	-	$\mathbf{9.7 \cdot 10^{-5}}$	$9.7 \cdot 10^{-1}$	-
(M5)	$1.2 \cdot 10^{-2}$	-	-	$1.46 \cdot 10^{3}$	-	-
(M6)	$9.4 \cdot 10^{-2}$	$1.1 \cdot 10^{-4}$	$\mathbf{9.4 \cdot 10^{-9}}$	$9.6 \cdot 10^{-1}$	$\mathbf{1 \cdot 10^{-9}}$	$\mathbf{4 \cdot 10^{-10}}$

RESULTS AND DISCUSSION.

When analytical aproximation $\widetilde{S}^*(\widetilde{\rho})$ of the solution (M1)-(M6) is found, the comparison of the quality of approximation is the main interset. Let us consider the quadratic error function $\mathcal{E}(\widetilde{S}^*)$:

$$\begin{aligned} \mathcal{E}_T &= \mathcal{E}_T(\widetilde{S}^*, \widetilde{S}_0, \phi, \widetilde{\nu}, \gamma, \sigma^2) \qquad (10) \\ &= \int_0^{\widetilde{S}_{0,v}} \int_0^{\phi_v} \int_0^{\widetilde{\nu}_v} \int_0^{\gamma_v} \int_0^{\sigma_v^2} \mathcal{E} d\widetilde{\nu} d\gamma d\sigma^2 d\phi d\widetilde{S}_0. \end{aligned}$$

where $\mathcal{E} = \int_0^1 \left(D^*(\widetilde{\rho}) \Delta \widetilde{S}^*(\widetilde{\rho}) - a^*(\widetilde{\rho}) \frac{\sigma^2 \widetilde{S}^*(\widetilde{\rho})}{1+\widetilde{S}^*(\widetilde{\rho})} \right)^2 d\rho$. Let us fix the parameters of bioreactor system by values having meaningful values of $\widetilde{S}_{0,v} = 10, \phi_v = 1, \widetilde{\nu}_v = 1, \gamma_v = 10, \sigma_v^2 = 10$. The error of approximations got from the considered models (M1) - (M6) is presented in Tables 1-2. The cases with '-' were not provided due to the complicated integration of loss expressions. In case of bold cases show the best approximation for each order.

Table 2: The errors $\mathcal{E}_T$ of models (M1) - (M6)

	0-order	1-order	2-order
(M1)	$3.7 \cdot 10^{4}$	$\mathbf{3.8 \cdot 10^{3}}$	$8.12 \cdot 10^{2}$
(M2)	$3.7 \cdot 10^{4}$	$\mathbf{3.8 \cdot 10^{3}}$	-
(M3)	$3.2 \cdot 10^{5}$	-	-.
(M4)	$3.3 \cdot 10^{3}$	$8.8 \cdot 10^{5}$	-
(M5)	$\mathbf{2.24 \cdot 10^{2}}$	-	-
(M6)	$7.9 \cdot 10^{4}$	$2.1 \cdot 10^{4}$	$\mathbf{2.7 \cdot 10^{2}}$

It can be seen in Tables 1-2 that the errors of different models vary by order of a few magnitudes. First it is worth to mention that the models (M3) and (M4) gives the significantly worse errors up to magnitude of order 10^5, which indicates that approximations are not useful. The model (M5) provides the best error at first order approximation; however, the complexity of this model is large and first or higher order approximation can not be easily evaluated. On the other hand, the model (M6) has the general form of approximation (9), k-th order approximations can be computed. By applying such linearization as in (M5) the equation is reduced to zero order kinetics. In such case the approximation works very well when concentration $\widetilde{S}$, but under larger concentrations the approximation is not working at all (see Table 1). Models (M1), (M2) and (M6) giving best approximation under the given fixed system parameters.

CONCLUSION

The non-linear DE based on Michaelis-Menten kinetics was solved using HPM with different initial linearizations. Different linearalization can significantly change the quality of the approximations (see Tables 1-2). The general form of $k-th$ order approximation was produced (9), which works well in a wide range of parameters.

Acknowledgements

The work was supported by the Research Council of Lithuania under Grant No. S-MIP-17-98.

REFERENCES

Baronas R.; Kulys J.; and Petkevičius L., 2018. *Modelling the enzyme catalysed substrate conversion in a microbioreactor acting in continuous flow mode. Nonlinear Anal Model Control*, 23, no. 3, 437–456.

Bender C.M. and Orszag S.A., 2013. *Advanced Mathematical Methods for Scientists and Engineers I: Asymptotic Methods and Perturbation Theory.* Springer Science & Business Media.

Eswari A. and Rajendran L., 2010. *Analytical solution of steady state current at a microdisk biosensor. Journal of Electroanalytical Chemistry*, **641**, no. 1, 35–44.

Fedorov A.A.; Berdnikov A.S.; and Kurochkin V.E., 2019. *The polymerase chain reaction model analyzed by the homotopy perturbation method. Journal of Mathematical Chemistry*, 57, no. 4, 971–985.

Filobello-Nino U.; Vazquez-Leal H.; Castaneda-Sheissa R.; Jimenez-Fernandez V.; Herrera-May A.; Huerta-Chua J.; Gil-Adalid L.; and Pretelin-Canela J., 2019. *An Analytical Approximate Solution for the Quasi-Steady State Michaelis-Menten Problem. Discrete Dynamics in Nature and Society*, 2019.

He J.H., 1999. *Homotopy perturbation technique. Computer methods in applied mechanics and engineering*, 178, no. 3-4, 257–262.

Saranya K.; Mohan V.; Kizek R.; Fernandez C.; and Rajendran L., 2018. *Unprecedented homotopy perturbation method for solving nonlinear equations in the enzymatic reaction of glucose in a spherical matrix. Bioprocess and biosystems engineering*, 41, no. 2, 281–294.

OBSERVABILITY OF VIRTUAL STOCHASTIC SENSORS: OBSERVABILITY TYPE-1

Pascal Krenckel, Claudia Krull, Graham Horton
Faculty of Computer Science
Otto-von-Guericke-University Magdeburg
Universitätsplatz 2, 39106 Magdeburg
Germany
E-mail: krenckel@ovgu.de, claudia.krull@ovgu.de

KEYWORDS
Virtual Stochastic Sensor, observability, augmented stochastic Petri nets, Proxel method, discrete stochastic models, discrete-time Markov chains

ABSTRACT

Virtual Stochastic Sensors (VSS) were developed for the analysis of partially observable discrete stochastic systems. In these systems, only some events produce observable output, which can be ambiguous. VSS enable behaviour reconstruction of augmented stochastic Petri nets (ASPN) based on system output protocols. The quality and utility of a VSS depends on how reliably it can reconstruct the internal state of a system from an observed output sequence. However, this result quality was not addressed in previous papers. In this paper, we quantify for the first time a measure of observability for VSS. This observability type-1 is defined as a measure of the ability to observe events. The experiments show that this measure makes it possible to predict the quality of the VSS in advance.

INTRODUCTION

In our daily life, stochastic processes can be found everywhere. For example, weather forecasts or durability of devices can be represented as stochastic processes. Virtual Stochastic Sensors (VSS) can reconstruct the behaviour of partially observable discrete stochastic systems (Krull et al. 2011). Like virtual sensors, VSS deduce not directly observable values from given measurements of other quantities. As opposed to a virtual sensor, the output and the input of a VSS is stochastic. In general, a VSS deduces the system behaviour from a given trace – a time series of physical sensor signals (Krull et al. 2011). Consider as an example a car rental agency with a single door sensor, where the quantity of interest is the number of customers waiting in the queue. Whenever a customer passes the door, this physical sensor may send a signal, which is the output used by the VSS to reconstruct the queue length. The theoretical utility of a VSS depends on how well the model represents the real world and the reliability of the output of the VSS.
We want to deal with the VSS performance and introduce a measure for the theoretical reliability of the reconstruction. Assessing the quality of a VSS was not addressed so far, but is interesting for two reasons. In a partially observable system, there is usually no ground truth available to compare the output of the VSS to. Therefore, quantifying the reliability of the behaviour reconstruction adds to the value of a VSS result. Secondly, a reliable measure of observability can determine if a certain sensor setup would lead to useful results before performing the actual sensor installation or behaviour reconstruction. This can also help to determine where to install sensors for best results.
In this paper, we introduce the concept of observability in the context of VSS. This paper is an extension of the master thesis of the first author (Krenckel 2019) and takes up the topic of observability type-1. The goal is to have an estimate of the quality of a VSS. The observability type-1 presented in this paper considers only the quantity of the output symbols and not the quality.

BACKGROUND

Observability in control theory

The term *observability* was introduced by Kálmán. In control engineering, it describes how well the internal state of a system can be determined by external measurements. Kálmán used the term to describe linear dynamic systems. In contrast to stochastic systems, linear dynamic systems are deterministic. If the starting state is known, the state variables can be determined at any discrete time. When the output vector has fewer dimensions than the input vector, it is not possible to determine the starting state from only one given data point. One can only learn about the initial state of the system if several measured values are considered over a time interval, to reconstruct the current system state (Kálmán 1970: Lunze 2002; Ham and Brown 1983).
"A state x is not observable in k steps if it is not distinguishable in the future in k steps from the zero state. It is not observable if it is not observable in k steps for all k." (Astolfi 2006). If the initial state is observable, the system is completely observable. In contrast to control engineering, the internal processes of the systems considered here are not deterministic, but stochastic. Therefore, internal states are not necessarily known at a later point in time, although the current system state may be known. Additionally, we assume that the initial state is known.

Petri-Nets

Carl Adam Petri developed Petri nets from finite-state machines. A Petri net is a bipartite graph with places and transitions as nodes. The arcs describe which places are the pre and post conditions of the transitions. Petri nets can model discrete, distributed systems as well as concurrency (Desek and Esparza 1995; Reisig 1985).
There are now many extensions of Petri nets. Augmented stochastic Petri nets associate a distribution function with each timed transition that maps the activation intervals. They also contain guard functions, arc multiplicities, inhibitor arcs and immediate transitions as well-known Petri net elements. The augmentation is that every transition can generate output symbols from a discrete set when firing (Buchholz 2012).
ASPN are used in this paper as user models for a VSS.

Proxel-Method

In 2002, Graham Horton invented the Proxel method based on supplementary variables (Cox 1955). This method is used to analyse systems with discrete states and continuous time, such as stochastic Petri nets or queuing systems. This method discretizes the continuous system time and explores the state space. A Proxel (probability element) is a tuple (p, S) where p is the probability of the proxel and S stands for the current state of the model, including the discrete system state and age of all currently active transitions, if needed. For each Proxel the follow-up Proxels are determined by calculating the probability that a state change is happening within the next time step, based on the ages of the corresponding state transitions (Horton 2002; Lazarova-Molnar and Horton 2003; Lazarova-Molnar and Horton 2004).
When analysing VSS with ASPN as user model, not only the current marking and transition ages are of interest. The definition of a Proxel is extended by including the previous system states. This enables behaviour reconstruction needed for VSS. We use a modified version of this behaviour reconstruction method for calculating the observability of a VSS (Buchholz 2012).

State-space explosion
The time and memory complexity of the Proxel method is exponential in the number of discrete time steps, which leads to a state space explosion, making computation impossible after only a few time steps. One method to reduce the state-space is merging Proxels with the same state and age vector, while cumulating their probability. This can however lead to difficulties when the previous system states are also coded in a Proxel, because the number of equal Proxels is greatly reduced (Horton 2002; Wickborn et al. 2006).

OBSERVABILITY OF VIRTUAL STOCHASTIC SENSORS

In this section, we will explain how observability of Virtual Stochastic Sensors in general is related to physical sensors. We will also give a first definition of the observability. Since the observability is only based on the quantity and not on the quality of the observations, we named it observability type-1. We will also clarify how the observability type-1 is related to the ability to reconstruct system states of VSS.

Physical and Logical Sensors

Physical sensors are responsible for the physical output of a system. Of course, the quality of the output depends on the quality and the quantity of the physical sensors. We differentiate between logical and physical sensors. A physical sensor is a real sensor in the real system. The logical sensor is the representation of a physical sensor in the model. Each set of physical sensors or a single physical sensor can be represented as one or more logical sensors in the model.
We assume that the physical sensors produce a digital output. Even if the sensors measure a continuous value, all results are quantized, to keep the system states discrete.
In the following, we explain how physical and logical sensors in general relate to system observability. The following rules can be used to superficially evaluate the effect of a sensor on the system observability.

1. *Activity level:* A high activity level means that the physical sensor is placed in a busy area of the system, making it likely to produce an output. The activity level only depends on the position of the sensor in the system. A higher activity level of a sensor is preferable, so that the sensor produces more observable outputs.
2. *Output granularity:* The granularity of the possible outputs of a sensor also has an influence on the quality of the reconstruction results of the VSS. If the processor activity of a computer has to be determined based on the processor fan volume, a fan with several speed levels can provide more information than a fan with only two speed levels. This physical sensor could be represented as several logical sensors in a system model. Due to a higher output granularity it is possible to distinguish different logical sensors. A higher output granularity of a sensor is preferable because it allows to distinguish different logical representations of the sensor.
3. *Output Ambiguity:* The term output ambiguity refers to the number of logical sensors with the same output. For example, a physical sensor may be represented at multiple locations in a model. A single light barrier at a door is represented by two logical sensors with the same output in the model, one at the entrance, one at the exit. A double light barrier can differentiate between an entering person and a leaving person. In this case, the two logical sensors have different outputs and are thus distinguishable. Less output ambiguity is preferable, to better distinguish between different events.
4. *Model Determinism:* A physical sensor in a part of the system with a high degree of determinism makes little sense. If even without a sensor it is already known what will happen, the sensor can only confirm what is already known. It is preferable to place a sensor in a less deterministic area of the system.

5. *Sensor Density:* The sensor density is closely related to model determinism. A high sensor density can lead to a high observability in a system part. If we know what happens in a system part, then the value of additional sensors there is very low. A sensor should therefore rather be placed in a system part with low sensor density. Exception: It has to be considered that the order of different sensor outputs can contain information on internal processes, e.g. for a double light barrier, two light barriers are located very close to each other. The sequence of the light barriers activity provides information about the direction of movement.

Based on these five points, possible positions of an additional physical sensor can be evaluated utilizing existing knowledge about the system. This evaluation catalogue gives a qualitative indication on the value of a specific sensor. In the next section, we will define observability type-1, which can quantify the impact of an additional sensor.

Observability Type-1

Observability type-1 represents a measure for the probability that an event can be observed under the condition that an event occurred, at a specific system time t: "If an event occurred, do I see it?" Mathematically this corresponds to:

$$O_1(t) = P(observed|event\ at\ time\ step\ t) \qquad (1)$$

Observability type-1 only depends on the percentage of observable events. Therefore, from the previous catalogue only point one (*1. Activity Level*) is covered.
Observability type-1 does not distinguish between different output symbols; *2. Output Granularity* and *3. Output Ambiguity* would require a distinction of the output symbols. In addition, two events with the same probability of occurrence have the same influence on the observability type-1. It does not matter what we already know about the system (*4. Model Determinism*), or how many sensors are already present in a part of the system (*5. Sensor Density*).

Computation
To calculate observability type-1, we use the Proxel method to generate a DTMC representing the complete reachable state space of the system. The states are expanded by the age variables and include the information if output was generated in the previous time step. In addition, the Proxel method assumes that only one event can occur in the system per time step. Therefore, observability type-1 is the sum of all Proxels probabilities created by an observable event divided by the sum of all Proxels created by an event (see Equation (2)).

$$O_1(t) = \frac{H_{observed}}{H_{event}} = \frac{\sum_{p \in proxel_{out_t}} P(p)}{\sum_{p \in proxel_{event_t}} P(p)} \qquad (2)$$

The past of the Proxels is not considered here. The corresponding model state space is a DTMC.
Each Proxel represents the state of the system, including time parameters and whether an output or event occurred. Since the past is not considered, it is also possible to merge Proxels over different time steps into a single Proxel. If the Proxels have the same state, all subsequent calculations are identical. As with every DTMC, the system time is not coded by the state itself, but by a parameter in the calculation.

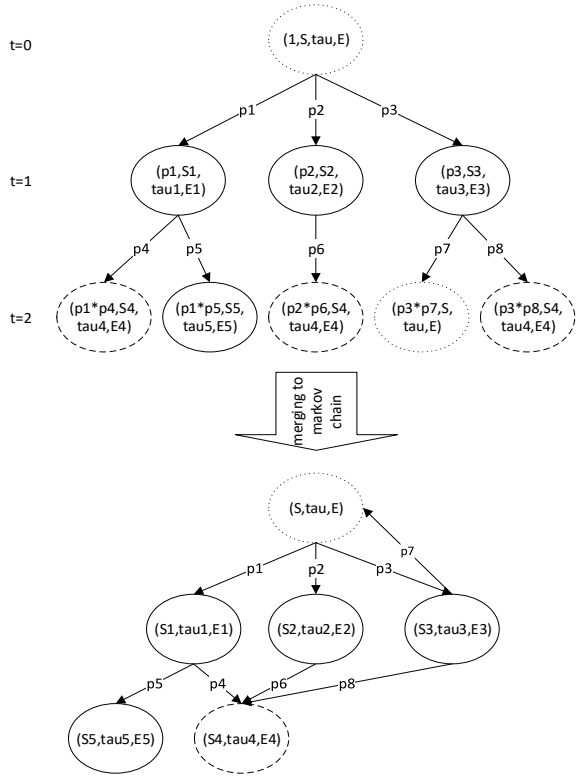


Figure 1: Conversion of State Space (Proxel Method) to DTMC. Dashed and Dotted Border indicate the same State.

Figure 1 shows the conversion of the already explored state space into a DTMC. Each Proxel is represented by a tuple (p, S, τ, O); p represents the total probability of the Proxel. S is the discrete system state (Petri-Net marking). τ contains the ages of the relevant transitions, and O indicates if an output was observed in the last time step. Since the computational structure corresponds to a DTMC, the result is equivalent to an HMM. The output O is observed with a probability of 1.
All similar states (same values for S, τ, O) can be represented by one state, since all subsequent calculations are identical.
The corresponding DTMC is not necessarily finite. With a queueing problem, a customer can join the queue at any time, even if the probability of doing so is approaching 0. The corresponding DTMC would have an infinite number of states. Therefore, it is necessary to limit the state space, e.g. by defining a maximum system time or a minimum probability of a state. However, this may cause the generated DTMC to contain absorbing states, which must be removed in order to allow solutions of the Markov chain beyond the simulated time. Since the removal of the absorbing states can result in new absorbing states, these must also be removed.
The DTMC generated using the Proxel algorithm can be used to calculate the observability type-1. The Markov chain must be solved up to time step t. The sum of the probabilities of the Proxels which were created by an event with output at

time step t divided by the sum of the probabilities of all Proxels which were created by an event at time step t results in the observability type-1 of time step t.

$$O_1(t) = \frac{\pi^0 * A^t * v_{out}}{\pi^0 * A^t * v_{event}} \quad (3)$$

Where $\pi^0 * A^t$ is the solution of the DTMC for time step t, v_{out} is a vector that indicates for each state whether it was created by an event with output and v_{event} is a vector that indicates for each state if it was created by an event.
Each state in the DTMC has a maximum of as many subsequent states as the number of possible events + 1 (nothing happens). Therefore, the transition matrix A is usually sparse. The computational effort for $\pi^0 * A^t$ is therefore in $O(n * t)$, where n is the number of states.
Combining Proxel algorithm and DTMC solution, observability type-1 can be calculated efficiently compared to the original Proxel method. In the next section, we examine how suitable observability type-1 is as a measure of the quality of a VSS.

EVALUATION – CASE STUDY

We use a model of a car rental agency also used in (Buchholz 2012) to evaluate our methods. The car rental agency is a queueing system. Figure 2 shows the ASPN of the system.

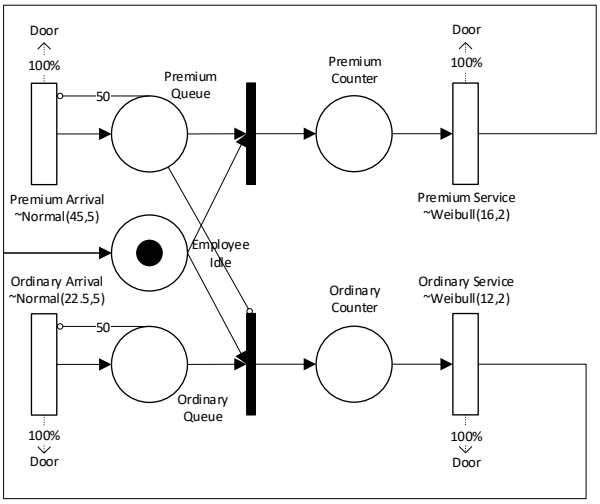


Figure 2: ASPN of a Car-Rental Agency (Base Model)

The system has an entrance through which customers enter or leave the building (*Premium Arrival* and *Ordinary Arrival*). The arrival times are normally distributed. At the door, there is a light barrier, which generates an output symbol every time a person passes the sensor (*Door*). There are two types of customer, premium and ordinary. Premium customers are served with priority and have a longer service time. The service times are Weibull distributed. In the model, each entry/exit of the normal/premium customers is represented by a corresponding timed transition (*Premium Service* and *Ordinary Service*). Both transitions emit the same output symbol when the service is finished and the customer leaves the building crossing the light barrier (*Door*). To test our method, we modified the system and created five distinct variations of this system setup. Each variation differs only in the output of the transitions.

1. *Base model (100%):* Each entry/exit produces the same output symbol (*Door*).
2. *Reduction of sensor reliability (70%):* To emulate a faulty sensor, each sensor only produces output with a certainty of 70%.
3. *Sensor reduction (in):* In this version, only incoming customers are detected, removing the output symbols at the service transitions.
4. *Sensor reduction (out):* In this version, output symbols are created only when a customer leaves the agency, removing the output symbols at the arrival transitions.
5. *Output uniqueness (I+O):* In this version of the model, the physical sensor is a double light barrier. It produces different outputs for incoming and exiting customers.

The quantity of interest is the queue length of premium and ordinary customers, as well as their sum. Figure 3 shows the observability type-1 for the different versions of the system.

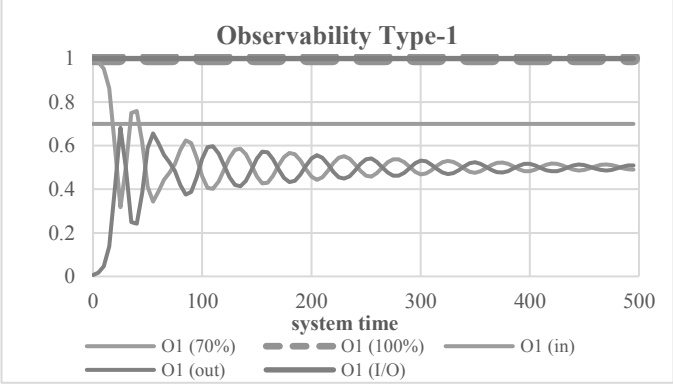


Figure 3: Calculated Observability Type-1 for the Different Systems

The computation results are not surprising. The base model (*100%*) and version 5 (*I+O*) have an observability of 1, since each transition reliably generates an output. Version 2 (*70%*) has an observability of 0.7. The observability of versions 3 (*in*) and 4 (*out*) is 0.5 because only half of the events can be observed. The order according to a decreasing observability type-1 is: version 5 (*I+O*) and version 1 (*100%*), version 2 (*70%*) and version 3/4 (*in/out*) last.
To evaluate the quality of our method, we measured the performance of the VSS to check if observability type-1 can assess the performance of the VSS with sufficient accuracy. The value of interest is the queue length from a given trace. Therefore, we use the expected deviation of the queue length as performance criterion. The higher the deviation, the greater the inaccuracy of the calculated queue lengths.
The performance results differ from the observability type-1. Figure 4 shows the standard deviation of the queue length of the VSS for the different model variations. Version 5 (*I+O*) scatters the least because the total queue can be calculated exactly from the trace. For this reason, the spread

of premium customers is just as large as the spread of ordinary customers. Surprisingly, the variations that only observe the input or the output have a smaller deviation than the base model, due to the missing output making the arrival and leave events distinguishable. The deviation of the total queue of version 3 (*in*) is greater than version 4 (*out*), since the deviation of the distribution functions of the *Arrival* transitions is greater than that of the *Service* transitions. However, the deviation of the premium queue is smaller due to the difference in the distribution functions between premium and normal customers is larger at the entrance. Both versions show that more information does not necessarily lead to better results, but that the results can even be worse. Version 2 (*70%*) also has less information than the base system, but there is a greater deviation here. This is expected since only the overall activity level was reduced.

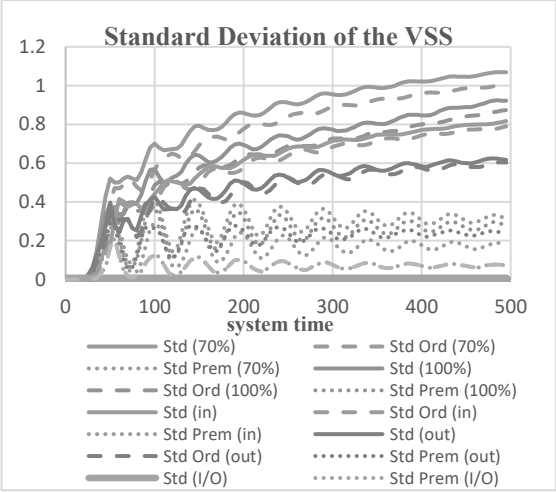


Figure 4: Expected Standard Deviation of the Total and Individual Queue Lengths of the Different Systems

The comparison of the observability type-1 values and the expected standard deviation of the different VSS shows that the observability type-1 does not consider the granularity and diversity of the outputs. The current definition considers only the activity rate, which has an influence on the quality of the sensor, but as variants 3 and 4 show, there are other factors as well. The difference between version 3 and 4 shows that the determinism of a single event also plays a role.

The experiment shows, that the quality of a VSS can be estimated prior to the actual deployment, assessing the possible gain of using VSS in a given real-world setting. We also showed, that observability type-1 can be used to compare different sensor setups, without installing them. However, the current definition is not detailed enough to reliably predict the performance of the VSS. More information on the sensor output and location needs to be considered to improve the estimate.

CONCLUSIONS AND FUTURE WORK

In this paper, we define observability in the context of Virtual Stochastic Sensors for the first time. We identified five parameters determining the gain of a physical sensor. This measure is evaluated and discussed using a case study. The five key parameters for the gain of a physical sensor are activity level, output granularity, output ambiguity, model determinism and sensor density. These parameters determine the added benefit of a physical sensor; i.e. how much more accurate the results of the VSS can become with the sensor. Observability type-1 only considers the activity level.

The observability type-1 is the average probability that an occurring event will also be observed. No distinction is made between different outputs, and the past is not considered. Thus, it is possible to calculate observability type-1 more efficiently than solving the complete VSS problem.

However, the case study shows that only considering the activity level is not sufficient to categorize the sensors according to their observability. This definition is therefore not yet usable in practice. Further research is needed to increase the quality of observability type-1. Our current research indicates that an additional differentiation of the output symbols can significantly improve the quality of the measure. How much the quality improves and where it can be applied needs to be investigated in the future.

REFERENCES

Astolfi, A. 2006. Systems and Control Theory An Introduction.

Buchholz, R. 2012. Conversive Hidden non-Markovian Models, Magdeburg.

Cox, D. R. 1955. „The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables“, Mathematical Proc. of the Cambridge Philosophical Society.

Desek, J. and J. Esparza. 1995. Free choice Petri nets, New York: Cambridge University Press.

Ham, F. M. and R. G. Brown. 1983. „Observability, Eigenvalues, and Kalman Filtering,“ in IEEE Transactions on Aerospace and Electronic Systems 2.

Horton, G. 2002. „A New Paradigm for the Numerical Simulation of Stochastic Petri Nets with General Firing Times“, in European Simulation Symposium, Dresden.

Kálmán, R. E. 1970. Lectures on Cntrollability and Observability, Stanford University.

Krenckel, P. 2019. Beobachtbarkeit von Virtuellen Stochastischen Sensoren, Otto-von-Guericke-Universität Magdeburg.

Krull, C.; R. Buchholz und G. Horton. 2011. „Virtual Stochastic Sensors: How to gain Insight into Partially Observable Discrete Stochastic Systems“, in The 30th IASTED International Conference on Modelling, Identification and Control.

Lazarova-Molnar, S. and G. Horton. 2003. Proxel-Based Simulation of Stochastic Petri Nets Containing Immediate Transitions.

Lazarova-Molnar, S. and G. Horton. 2004. Proxel-Based Simulation of Stochastic Petri Nets, SCS European Publishing House.

Lunze, J. 2002. Regelungstechnik 2: Mehrgrößensysteme, Digitale Regelung, Springer-Verlag Berlin Heidelberg GmbH.

Reisig, W. 1985. Petri Nets: An Introduction, Springer-Verlag.

Wickborn, F.; C. Isensee; T. Simon; S. Lazarova-Molnar and G. Horton. 2006. A New Approach for Computing Conditional Probabilities of General Stochastic Processes.

SIMULATION OPTIMIZATION

ONE DIFFICULT OPTIMIZATION PROBLEM
AND
ONE IMPOSSIBLE ONE PLUS GOAL PROGRAMMING

William Conley
Austin Cofrin School of Business
University of Wisconsin at Green Bay
480B Wood Hall
Green Bay, Wisconsin 54311
U.S.A.
E-mail: Conleyw@uwgb.edu

KEYWORDS
Complex equations, nonlinear optimization, goal programming.

ABSTRACT

Presented here is a large nonlinear system of equations where an exact solution is desired or required. Multi stage Monte Carlo optimization (MSMCO) does succeed in solving this system of equations with no error. The technique is explained and illustrated.

However, there are many cases in business and other technical fields where there are so many nonlinear goals (or equations) that the organization has a lot of difficulty in solving the entire system (meeting all the goals). Hence, the term or field of goal programming arose. The idea being to work out a useful approximate solution when an exact solution is unattainable.

INTRODUCTION

Consider the system of equations:

$x_1 + x_2 + x_3 = 19$
$5x_1 + 2x_2 + x_3 = 58$
$x_1 + 9x_2 + 6x_3 = 95$

This system can be solved by minimizing:

$g(x_1,x_2,x_3) = |x_1 + x_2 + x_3 - 19| + |5x_1 + 2x_2 + x_3 - 58| + |x_1 + 9x_2 + 6x_3 - 95|$

down to zero using statistical optimization simulation techniques.

This problem also could be easily solved using the well known linear algebra techniques (substitution, cancellation, etc.).

However, let us use the statistical optimization (MSMCO) simulation approach on the following more difficult nonlinear system of equations.

THE LARGE NONLINEAR SYSTEM

The problem is to find an all whole number solution to the nonlinear system of equations

$x_i^3 + x_{111-i} \cdot x_{110} - i = C_i$

for i = 1, 2, 3, . . . 109

and

$x_{110}^3 + x_{37} \cdot x_{65} = C_{110}$

in the range of $1 \le x_i \le 70$ and all whole numbers for i=1,2,3, . . . 110 where the right-hand side constants C_i for i = 1, 2, . . . 110 follow here in the pattern

C_1	C_2	C_3	C_4	C_5
C_6	C_7	C_8	C_9	C_{10}
"	"	"	"	"
"	"	"	"	"
"	"	"	"	"
"	"	"	"	"
"	"	"	"	"
C_{106}	C_{107}	C_{108}	C_{109}	C_{110}

Table 1: Right Hand Side of Equations Constants

68938	9317	143128	178181	3622
6507	301168	8480	31711	4652
219864	330189	1916	99550	87726
345294	29479	205915	1105	55154
152	60855	332093	304683	4087
330307	17082	93193	100416	11619
252631	251035	43421	316	869
1496	43883	166417	1741	8637

3473	18024	343512	151053	85456
296	364	141658	150732	344696
50877	749	3681	27665	183
140695	250134	195133	602	35358
152587	45631	27208	1024	776
41636	265854	2332	175850	117829
2437	661	76013	13894	238338
1748	9401	19781	58841	315629
343874	87254	104948	31516	195319
343201	180239	264835	13902	292
104165	1043	302533	52753	241408
70945	158660	28794	179756	328989
216248	33388	4715	20889	15859
185921	94037	176708	862	5193

THE SOLUTION APPROACH

Denoting the left-hand side of equation i as L_i in our 110X110 system and the right hand side of equation i as C_i (the aforementioned constants) the MSMCO statistical stimulation optimization program tries to minimize

$$f(x_1,x_2,x_3 \ldots x_{110}) = \sum_{i=1}^{110} |L_i - C_i| \text{ down to zero keeping in mind}$$

that the x_i variable values must be whole numbers in the range of 1 to 70. Therefore, there are 70^{110} feasible solutions (about $9.1365956 \times 10^{192}$ possible answers).

The program drew 50,000 real value sample feasible solutions and stored the best answer (lowest total equation error). Then centered about this "best answer so far" in a smaller search region (all 110 dimensions reduced by a factor of 1.411 or the square root of 2) another sample of 50,000 more feasible solutions is looked at and the best one stored. Then in stage 3 centered about this new "best answer so far" another 50,000 feasible solutions are looked at in a similarly reduced ($\sqrt{2}$) search region and the best one is stored, etc. This process is done 25 times and it just rockets across the feasible solution space of 110 dimensions to a near approximation of the solution.

This answer is stored. Then the whole process is repeated for another 25 stages drawing 25x50,000 more feasible solution and its best answer stored also.

Then the variable values from the first 25x50,000 samples "best answer" are compared with the variable values from the second set of 25x50,000 feasible solutions and any variables that have these two values within one tenth of each other are averaged and rounded to the nearest whole number between 1 and 70. Then these values are fixed for the rest of the simulation.

Then this whole process is repeated twice more (with those fixed values) and random samples of the rest of the region for 25x50,000 more samples. Then any of the variables in the two new best answers that have their two values also less than .1 are averaged and rounded to the nearest whole number and fixed for the rest of the simulation.

Then two more sets of 25x50,000 sample answers are done and compared and any two more variable values that are less than .1 apart are averaged and rounded to the nearest whole number and fixed.

Then lastly, one more sample of 25 stages times 50,000 = 1,250,000 samples is looked at and the whole answer emerges once the remaining free floating x_i values are rounded to the nearest whole numbers.

The answer follows in the next section. It checks exactly in all the equations.

THE SOLUTION

The solution is in the pattern:

x_1=41	x_2=21	x_3=51	x_4=56	x_5=13
x_6=18	x_7=67	x_8=20	x_9=31	x_{10}=8
x_{11}=60	x_{12}=69	x_{13}=26	x_{14}=46	x_{15}=44
x_{16}=70	x_{17}=30	x_{18}=59	x_{19}=9	x_{20}=38
x_{21}=2	x_{22}=39	x_{23}=69	x_{24}=67	x_{25}=3
x_{26}=69	x_{27}=25	x_{28}=45	x_{29}=46	x_{30}=19
x_{31}=63	x_{32}=63	x_{33}=35	x_{34}=4	x_{35}=5
x_{36}=2	x_{37}=35	x_{38}=55	x_{39}=12	x_{40}=20
x_{41}=9	x_{42}=26	x_{43}=70	x_{44}=53	x_{45}=44
x_{46}=6	x_{47}=4	x_{48}=52	x_{49}=53	x_{50}=70
x_{51}=37	x_{52}=7	x_{53}=3	x_{54}=29	x_{55}=3
x_{56}=52	x_{57}=63	x_{58}=58	x_{59}=7	x_{60}=32
x_{61}=53	x_{62}=35	x_{63}=30	x_{64}=10	x_{65}=8
x_{66}=34	x_{67}=64	x_{68}=8	x_{69}=56	x_{70}=49
x_{71}=13	x_{72}=1	x_{73}=42	x_{74}=24	x_{75}=62
x_{76}=12	x_{77}=21	x_{78}=26	x_{79}=38	x_{80}=68
x_{81}=70	x_{82}=44	x_{83}=47	x_{84}=31	x_{85}=58
x_{86}=70	x_{87}=56	x_{88}=64	x_{89}=24	x_{90}=6
x_{91}=47	x_{92}=8	x_{93}=67	x_{94}=37	x_{95}=62
x_{96}=41	x_{97}=54	x_{98}=30	x_{99}=56	x_{100}=69
x_{101}=60	x_{102}=32	x_{103}=15	x_{104}=27	x_{105}=25
x_{106}=57	x_{107}=45	x_{108}=56	x_{109}=1	x_{110}=17

Figure 1 shows a two-dimensional projection of the seven 25x50,000=1,250,000 (times7) or 8,750,000 sample solution tries going across the feasible solution space looking for and finding the solution.

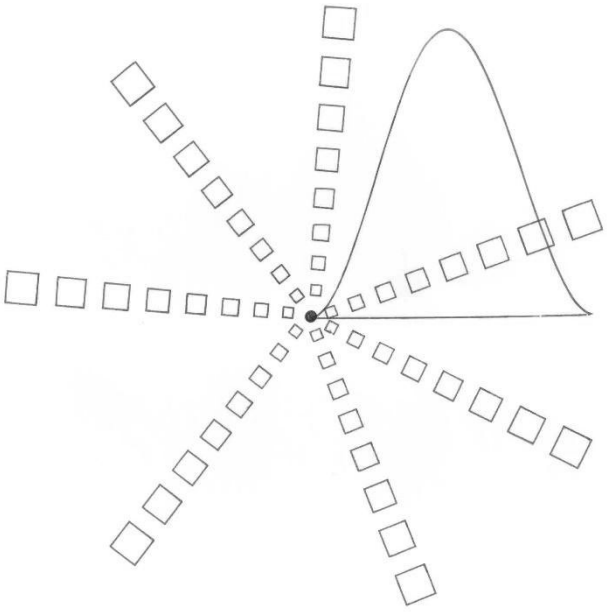


Figure 1: Statistical Optimization MSMCO Looking for and Finding the True Optimal Solution

DISCUSSION OF THE ALGORITHM

Note that even though the correct solution was produced the element of risk was present in the simulation. Observing that x_1=41is in the true optimal solution, what if it was unfortunately pinned down at 42 somewhere in the simulation? Then, of course, the "answer" that emerged at the conclusion would be wrong. That is why it is important to draw a lot of samples. Would it be better to use a mode of 3 and only pin variable values down if three tries in a row gave extremely close values (about the same or exactly the same values)? The author has experimented with modes of 3 through 8 on a variety of problems that are not systems of equations (Conley 1991a) and (Conley 1991b). However, as computers get faster and faster lower modes are less risky. With slower computers, higher modes may be more attractive.

Also, the range of closeness was fixed at .1 for this problem (after some experimenting). Would a closeness range of .03 or .09 be better? Possibly. What really matters is that some sort of statistical averaging of the coordinates (of the best answers so far) take place. Averages of sample means, medians, modes, harmonic means, geometric means or even a mean of all five of the means should work. Also, instead of a function of the range for closeness (.1 in this case) functions of the standard deviation or mean absolute deviations will also work.

Statistical optimization (MSMCO) needs powerful computers to work well. However, these seem to be fairly inexpensive and omnipresent in our new century. The only real issue is local optimals and even they tend to be useful when they pile up around the true optimal, which frequently happens. Then they can be exploited statistically to find the true optimal or a better useful approximation.

It can not be emphasized enough that the tremendous increase in speed and capacity of computers opens up the whole field of statistical optimization. (Conley 1989) used a mode=7 approach and only drawing a sample of 1000 feasible solutions in a two stage MSMCO approach (repeated seven times) to find the shortest route connecting the 88 cities in Ohio, USA (that are the county seats of government) in a shortest route using great circle distances (from their latitudes and longitudes). Then pinning down sub-routes that occurred seven times in a row and repeating until the "best" route emerges. This is a totally different problem (not a system of equations) from an era when computer time was expensive.

Therefore, drawing 8,750,000 sample answers as was done here was out of the question. However, the modal averaging idea was the same (adjusted to Mode=7 because the risk was higher in 1989 with less samples drawn). (Conley 1988) worked on a routing problem to control costs by saving airline miles from (Eilon, et al 1971) using modal averaging with limited computing power (and Pythagorean distance measures) and found the true optimal solution.

GOAL PROGRAMMING

The term "goal programming" could mean different things to lots of people. However, in the field of multivariate optimization (or Mathematical Programming) it can mean a system of equations (linear or nonlinear) where there are the same numbers of, or more equations than variables, which leads one to think there may be no exact solution.

It is desired to try to find an all whole numbers solution of the nine equation and nine variable system of equations in the range $1 \le x_i \le 300$ for i=1,2,3, . . 9.

$x_1 + x_3 + x_5 + x_7 + x_9 = 553$ (1)
$x_2 + x_4 + x_6 + x_7 + x_9 = 646$ (2)
$x_3 + x_5 + x_8 + x_9 = 395$ (3)
$x_2^2 + x_3^2 + x_6^2 + x_7x_8x_9 = 286{,}577$ (4)
$x_1^2 + x_6x_7 + x_9 = 60{,}881$ (5)
$x_1^3 + x_5^2 + x_6 + x_8 = 55{,}608$ (6)
$x_4^3 + x_7 + x_9 = 14{,}348{,}996$ (7)
$x_6^2 + x_7 + x_8^2 + x_9 = 61{,}443$ (8)
$x_2\, x_6\, x_8 + x_9 = 2{,}618{,}082$ (9)

Because the feasible solution space is relatively small 300^9 possible solutions, we will use regular multi stage Monte Carlo optimization (Statistical Optimization) without averaging.

However, because some of the constants (the right hand sides of the equations) are quite small and some of them are large we will use weights centered about one million so that our computer program treats them all equally. This makes for a smoother simulation.

Therefore. Letting the left hand side of each equation be L_i and the right hand side be C_i and the weights be $W_i=1000000/C_i$ the problem is transformed to minimize

$$f(x_1,x_2,x_3 \ldots x_9) = \sum_{i=1}^{9} |L_i - C_i| W_i \text{ subject to } 1 \le x_i \le 300 \text{ and}$$

whole numbers and remember to take the weights (W_i's) off when printing the equation errors at the end of the program.

The program uses 15 stages of ever decreasing in size and free floating search of 5,000,000 feasible solutions at each stage.

The "solution" produced was:

$x_1 = 218$
$x_2 = 99$
$x_3 = 157$
$x_4 = 243$
$x_5 = 88$
$x_6 = 215$
$x_7 = 62$
$x_8 = 123$
$x_9 = 27$

with an error of one unit in equation 1 and two units in equation 6 and no errors in the other seven equations.

Even though an exact solution was not found it still might be useful in a management goal programming setting.

Please see Figure 2 for a partial geometric and statistical representation of statistical optimization (MSMCO) simulation.

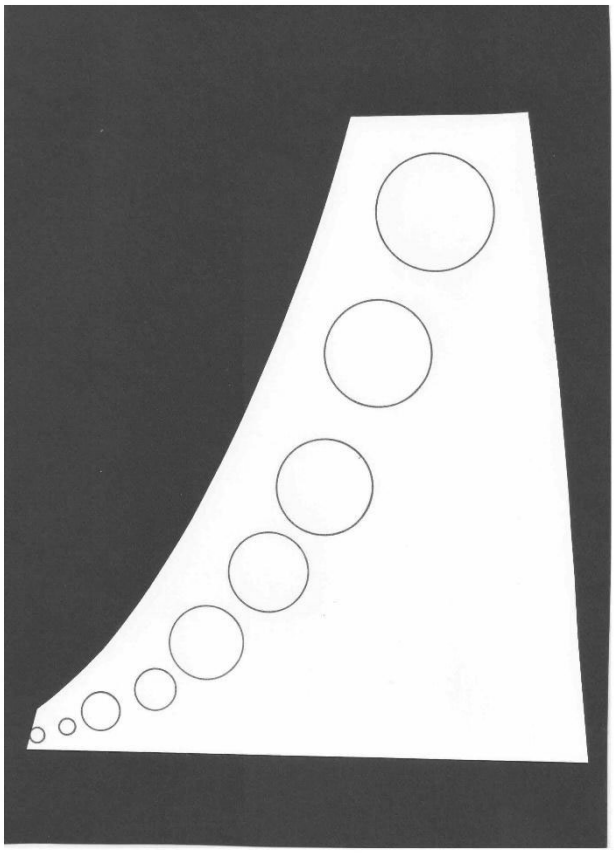


Figure 2: MSMCO Looking for the Minimum Error Solution Region

However, frequently in economics, big business and science there are so many goals that an "approximate solution" of some type that comes close to achieving many of the goals can be very useful and desirable.

Earlier it was mentioned in the previous 110x110 system that was solved with statistical averaging, that if one of the variables, say x_1 was pinned down incorrectly at 42 when it should be 41, then of course the whole answer is "wrong". However, if the other 109 values are correct, this could still be a very useful approximate solution out in the "real world" of science, economics and business, etc.

Therefore, if a corporation had offices and production plants and sales forces in ten states, if they were all looked at together in one corporate planning and operations, then there would be many goals that are all interrelated. There might be 10 revenue and 10 cost goals

(hence 20 equations) for the corporation in the 10 states. There might be additional sales goals, market share goals and overall corporate profit goals and minimum pollution goals, etc.

This could lead to massive systems of equations which would probably be unsolvable and/or have no exact solution. However, top management would be very interested in an "approximate solution" that say represented an improvement over the previous year of 1 or 2 or 5 percent.

Therefore, if these problems are not solvable by traditional methods, statistical optimization (multi stage Monte Carlo optimization with averaging) is always available to survey the feasible solution space of all the answers and try to close in on the optimal solution or a hopefully useful approximation.

Other examples might be a company or organization that operates in or is concerned with all 46 countries in Europe. Then as the goals are set up perhaps hundreds of equations and variables could be involved. Statistical optimization with averaging is always available in our computer age for the quantitative part of the planning.

Any plans for pollution control probably would be international with many variables and equations to be solved. Some of the pollution from one country or state could end up in other countries or states that are "down wind" or "down river" from the source. This is true in all six of our inhabited continents.

Again, statistical optimization with averaging (hooked up to a powerful computer) could at least come up with an approximate solution to these massive optimization problems. Optimization is so important in engineering too.

Additionally, we all know that many of these problems involve factors that are not quantitative. However, statistical optimization could still help solve the technical part, which would then be available to the "decision makers" who also have to consider the nonquantitative factors before taking action.

CONCLUSION

Presented here was a statistical optimization simulation approach to solving a 110 variable nonlinear system of 110 equations for the solution using the principles of statistics and computer simulation in our modern times where there is enough computer power available to draw 8,750,000 sample feasible solutions and keep track of them statistically until they point the way to and ultimately find the solution in about two minutes of computer run time.

Good reviews of statistics can be found in (McClave, Benson, Sincich 2001), (Anderson, Sweeney, Williams 1999), (Anderson 2003), (Hayter 2002) and (Keller and Warrack 2003). Note also that (Conley 2017) mentioned the modal and statistical averaging but did not use it in solving the four nonlinear systems of equations presented because they ranged in size from five to twenty-five variables and it was not necessary in those problems. Goal programming was discussed and its potential because it frequently leads to very complex nonlinear systems of equations. Additionally, a small hypothetical example of nine equations and nine variables system of equations with no exact solution was presented and "solved" for the true minimum total error of three in the goal program section, that deals with problems without an exact (no error) solution.

REFERENCES

-Anderson, D. R., Sweeney, D. J., Williams, T. A. 1999. *Statistics for Business and Economics*. 7th Edition. Southwestern College Publishing, Cincinnati, Ohio.

-Anderson, T. W. 2003. *Multivariate Statistical Analysis*. 3rd Edition. Wiley and Sons, New York.

-Black, K. 2014. *Business Statistics for Contemporary Decision Making*. 8th Edition. John Wiley and Sons, New York.

-Conley, W. C. 2017. "Statistical Optimization Applied to a Variety of Nonlinear Systems of Equations." In *Proceedings of the 2017 European Simulation and Modelling Conference*. Ostend, Belgium. ESM207-ETI, 54-59.

-Conley, W. C. 1988. "Controlling Costs by Finding the Shortest Route." *Cost Engineering Journal*. November, Vol 30, No. 11, EUROSIS, 20-24.

-Conley, W. C. 1989. "Finding the Shortest Routes in Ohio." *Proceedings of the Twentieth Annual Meeting of the Midwest Decision Science Institute*. Miami University, Ohio, 219-221.

-Conley, W. C. 1991. "Multi Stage Monte Carlo Optimization Applied to a Two Hundred Point Traveling Salesman Problem." *Proceedings of the 1991 Summer Computer Simulation Conference* (Baltimore). The Society for Computer Simulation, SCS, San Diego, 145-151.

-Conley, W. C. 1991. "Programming an Automated Punch or Drill." *International Journal of Systems Sciences*, Vol. 22, No. 11, 2039-2056.

-Eilon, S., Watson, Gandy C. and Christofides, N. 1971. *Distribution Management, Mathematical Modeling and Practical Analysis*. Griffon, London, England.

-Hayter, A. J. 2002. *Probability and Statistics for Engineers and Scientists*. 2nd Edition. Duxbury Press, Pacific Grove, CA.
-Keller, G. and Warrack, B. 2003. *Statistics for Management and Economics*. 6h Edition. Thompson Brooks/Cole, Pacific Grove, CA.
-McClave, J. T., Benson, P.G., Sincich, T. 2001. *Statistics for Business and Economics*. 8th Edition. Prentice Hall, Inc. Upper Saddle River, New Jersey.

BIOGRAPHY

WILLIAM CONLEY received a B.A. in mathematics (with honors) from Albion College in 1970, an M.A. in theoretical mathematics from Western Michigan University in 1971, a M.Sc. in statistics in 1973 and a Ph.D. in mathematics-computer statistics from the University of Windsor in 1976. He has taught mathematics, statistics, and computer programming in universities for over 30 years. He is currently a professor emeritus of Business Administration and Statistics at the University of Wisconsin at Green Bay. The developer of multi stage Monte Carlo optimization and the CTSP multivariate correlation statistics, he is the author of five books and more than 230 publications world-wide. He is a member of the American Chemical Society, a fellow in the Institution of Electronic and Telecommunication Engineers and a senior member of the Society for Computer Simulation. He gave a speech and presentation on his research at the 7th International Conference on Genetic Algorithms in July 1997 in the United States.

STATISTICAL OPTIMIZATION WITH BALANCING WEIGHTS AND A MODE OF THREE

William Conley
Austin Cofrin School of Business
University of Wisconsin at Green Bay
480B Wood Hall
Green Bay, Wisconsin 54311
U.S.A.
E-mail: Conleyw@uwgb.edu

KEYWORDS
Optimization, controlling risks or errors, nonlinear systems, complex equations, Monte Carlo.

ABSTRACT

Two systems of nonlinear equations are presented and solved using statistical optimization or multi stage Monte Carlo simulation optimization. The equations have balancing weights attached to them to overcome the problem that can be caused when one or more equation constants is so small compared to the other equation's constants (right hand sides of the equations) that the simulation almost ignores it. That can lead to incorrect answers at the end of the simulation.

The first system has all whole number exponents and the optimal solution (no error) is produced. The second system has fractional powers (exponents) throughout and even though little or no equation errors are reported in the printout there may be slight equation errors due to rounding and other factors. This will be discussed along with the odds, that even though there are rounding errors the probability is still there that the true optimal was found with simulation applied to optimization using a modal averaging scheme of three.

INTRODUCTION

The mode in the world of statistics is the most frequently occurring number. The data set 1, 2, 2, 2, 3 and 3 has a mode of 2. It is a measure of central tendency or average. (Conley 1984) on pages 147-159 of his book presented a 12 variable nonlinear optimization that multi stage Monte Carlo optimization (MSMCO) tried to solve one hundred times on a university's main frame computer. It produced one hundred approximate answers that were all incorrect. However, it was noted on page 159 of the book that the coordinate-wise mode and median produced the true optimal solution. Also, the coordinate-wise mean rounded to the nearest whole number was the optimal solution.

With our 21st century, more powerful, relatively inexpensive and omnipresent computers today, engineers, scientists and business people can solve a large variety of optimization problems using this averaging approach. Let's try a mode of three (drawing larger samples than were possible last century) with today's powerful computers on two challenging nonlinear systems of equations. Also, "balancing weights" will be attached to the systems of equations to smooth the simulations across the sampling distributions of the feasible solution spaces.

A NOTE ON BALANCING WEIGHTS

Consider the system of equations

$$\sum_{i=1}^{7} x_i = 5000 \quad (1)$$

$$\sum_{i=1}^{7} x_i^2 = 500{,}000 \quad (2)$$

$$\sum_{i=1}^{7} x_i^3 = 50{,}000{,}000 \quad (3)$$

Using statistical optimization (MSMCO) to minimize

$$f(x_1,x_2,\ldots x_7) = \left|\sum_{i=1}^{7} x_i - 5000\right| + \left|\sum_{i=1}^{7} x_i^2 - 500{,}000\right| +$$

$$\left|\sum_{i=1}^{7} x_i^3 - 50{,}000{,}000\right|$$ difficulties could arise because the

right hand side constant in equation (1) is only 5000. That is small compared with the constants 500,000 and 50,000,000 in equations (2) and (3). Therefore, to make them more even we use "balancing weights" of $w_1=100$ for equation (1) and $w_3=.01$ for equation (3) and try to minimize

$$g(x_1,x_2,\ldots x_7) = \left|\sum_{i=1}^{7} x_i - 5000\right| \cdot 100 + \left|\sum_{i=1}^{7} x_i^2 - 500{,}000\right|$$

$$+ .01\left|\sum_{i=1}^{7} x_i^3 - 50{,}000{,}000\right|$$

Also remember that it is a good idea to remove the balancing weights later in the simulation before printing any error terms so that the error terms will be more accurate.

PROBLEM ONE

Consider the system of equations

$$x_i^4 + x_{i+1}^3 + x_{i+2} x_{i+3} = C_i$$

for $i = 1,2,3\ldots.37$

and

$$x_{38}^3 + x_1 x_2 x_{34} = C_{38}$$

$$x_{39}^4 + x_{25} x_{37} + x_{31} x_{36} = C_{39}$$

$$x_{40}^3 + x_3 x_4 + x_{38} x_{39} = C_{40}$$

and all x_i's are whole numbers from 1, 2, … to 50 where the C_i right hand constants are given here in the pattern

c_1	c_2	c_3	c_4	c_5
c_6	c_7	c_8	c_9	c_{10}
c_{11}	c_{12}	c_{13}	c_{14}	c_{15}
c_{16}	c_{17}	c_{18}	c_{19}	c_{20}
c_{21}	c_{22}	c_{23}	c_{24}	c_{25}
c_{26}	c_{27}	c_{28}	c_{29}	c_{30}
c_{31}	c_{32}	c_{33}	c_{34}	c_{35}
c_{36}	c_{37}	c_{38}	c_{39}	c_{40}

619905	85064	113643	5325107	408898
526611	2830716	157652	3113632	81353
3511935	4171016	2944910	5793651	936073
6302103	1900650	832785	6324196	3113478
21513	7137	3788	95749	4186349
3750364	112257	410364	58311	1245450
2387599	3112451	5169	131381	4165729
2658440	4491648	20488	4481086	1952

Therefore, the attempt is made to minimize

$$g(x_1,x_2,\ldots x_{40}) = \sum_{i=1}^{37} w_i\left|x_i^4 + x_{i+1}^3 + x_{i+2} x_{1+3} - C_i\right|$$

$$+ \left|x_{38}^3 + x_1 x_2 x_{34} - C_{38}\right|W_{38} + \left|x_{39}^4 + x_{25} x_{37} + x_{31} x_{36} - C_{39}\right|W_{39}$$

$$+ \left|x_{40}^3 + x_3 x_4 + x_{38} x_{39} - C_{40}\right|W_{40}$$

for $1 \le x_i \le 50$ and all whole numbers and $W_i = 10000/C_i$ for all i=1, 2, 3, 4, 5, … 40.

The simulation plan is to have 25 stages of 3 MSMCO simulations (we are using a mode of three) using real values arithmetic. Each of those 3 MSMCO simulations will draw 50,000 feasible solutions 16 times as 16 ever decreasing in size and free floating "rectangles" try to cross the sampling distribution of $g(x_1 \ldots x_{40})$ to an approximate solution. After the first of the 25 stages is done three times, the 3 approximate answers are compared coordinate wise and any variable where these 3 values are within .025 of each other are averaged (the sample mean) and rounded to the nearest whole number. That x_i value is then pinned down for the remining 24 sets of 3 solution tries. Then in an outer stage 2 this process is repeated with more x_i's being pinned down.

This outer stage process with subsequent modal averaging of three pinning downs is done 23 more times and the whole number answer emerges (as the weights W_i's are removed before the error terms are printed). Please see Figures 1 and 2 for a partial geometric and statistical representation of MSMCO (multi stage Monte Carlo optimization) at work.

The solution is in the pattern

x_1	x_2	x_3	x_4	x_5
x_6	x_7	x_8	x_9	x_{10}
x_{11}	x_{12}	x_{13}	x_{14}	x_{15}
x_{16}	x_{17}	x_{18}	x_{19}	x_{20}
x_{21}	x_{22}	x_{23}	x_{24}	x_{25}
x_{26}	x_{27}	x_{28}	x_{29}	x_{30}
x_{31}	x_{32}	x_{33}	x_{34}	x_{35}
x_{36}	x_{37}	x_{38}	x_{39}	x_{40}

28	17	7	48	25
26	41	17	42	1
43	45	41	49	30
50	37	29	50	42
12	9	6	8	45
44	12	45	12	33
39	42	5	14	45
40	46	24	46	8

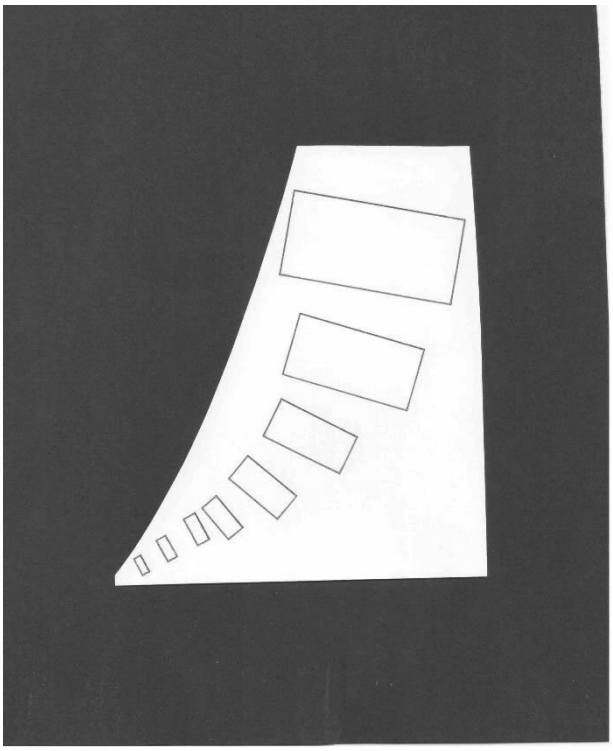


Figure 1. Statistical Optimization Heading Toward the Minimum Error

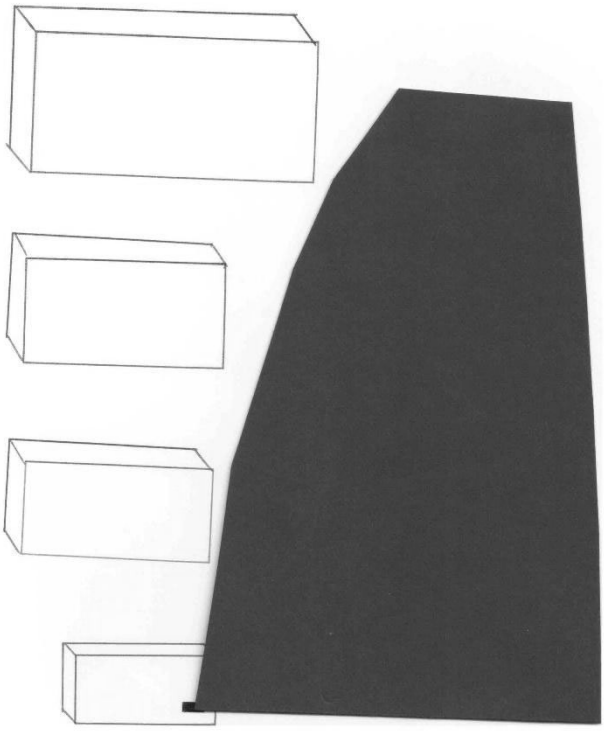


Figure 2. Looking for the Solution in a Three Variable Problem

The equation error terms are in the same pattern and they are all zero.

0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00

The 25x3x16x50,000=60,000,000 sample feasible solutions are a lot of possible solutions to look at and average.

However, thousands or billions of "local optimals" could be "hiding" in forty-dimensional space. Therefore, it is good to be cautious with the probability and statistics and geometry calculations going on in this simulation to manage the risk. Risk management is important to financial planners, and insurance actuaries and it is also important in all manner of simulation problems in many different fields.

(Anderson 2003), (Black 2014) and (Anderson, Sweeney, Williams 1999) offer good reviews and advanced theories in the field of statistics. Let us look at another example.

PROBLEM TWO: FRACTIONAL POWERS

Looking at the following system of nonlinear equations with fractional exponents that require all whole number solutions such that the x's all are in the 1 or 100 range or $1 \leq x_i \leq 100$ for i=1, 2, 3, … 35, let us try simulation again. This time equation balancing weights of $10000/C_i$ are used as in the previous problem (and they are removed before printing the final error terms).

Studying the 35 constants (the right hand sides) and all of the fractional powers it seems unlikely that an all whole number solution would produce no errors in the equations. However, in theory, it is possible even though the true optimal (smallest total error) probably will have some small error terms.

$$x_i^{3.4} + x_{i+1}^{2.1} + x_{i+2}^{.19} x_{i+3}^{.25} = C_i$$

for i = 1, 2 . . . 32

$$x_{33}^{3.7} + x_1^{.2} x_3^{.1} + x_7^{.3} x_9^{.4} = C_{33}$$

$$x_{34}^{3.5} + x_2^{.1} x_3^{.3} + x_{11}^{.2} x_{15}^{.5} = C_{34}$$

$$x_{35}^{3.9} + x_1^{.2}\, x_4^{.4} + x_{27}^{.5}\, x_{20}^{.4} = C_{35}$$

where the C_i right hand constants are given here in the pattern

c_1	c_2	c_3	c_4	c_5
c_6	c_7	c_8	c_9	c_{10}
c_{11}	c_{12}	c_{13}	c_{14}	c_{15}
c_{16}	c_{17}	c_{18}	c_{19}	c_{20}
c_{21}	c_{22}	c_{23}	c_{24}	c_{25}
C_{26}	C_{27}	C_{28}	C_{29}	c_{30}
c_{31}	c_{32}	c_{33}	c_{34}	c_{35}

500176.1	6324127.5	5497324.0	1111492.4	35278.3
4586410.5	1969751.9	31491.0	18905.5	5303515.5
604996.09	1795513.9	2374373.8	203126.8	1890357.5
5115086.5	331765.9	118456.1	57240.4	2065271.9
10478.2	776518.0	5537.5	1111225.3	12344.1
1388479.5	943387.0	3090545.0	2268423.0	184179.8
1467657.9	2607298.3	17009742.0	1314214.3	3906346.3

This system is transformed to minimize

$$g(x_1,x_2,\ldots x_{35}) = \sum_{i=1}^{32} |x_i^{3.4} + x_{i+1}^{2.1} + x_{i+2}^{.19}\, x_{i+3}^{.25} - C_i| \cdot W_i$$

$$+ |x_{33}^{3.7} + x_1^{.2}\, x_3^{.1} + x_7^{.3}\, x_9^{.4} - C_{33}|W_{33}$$

$$+ |x_{34}^{3.5} + x_2^{.1}\, x_3^{.3} + x_{11}^{.2}\, x_{15}^{.5} - C_{34}|W_{34}$$

$$+ x_{35}^{3.9} + x_1^{.2}\, x_4^{.4} + x_{27}^{.5}\, x_{20}^{.4} - C_{35}|W_{35}$$

where the W_i's are $10000/C_i$ and $1 \le x_i \le 100$ and whole numbers for i=1, 2, ... 35 and the W_i's for i=1, 2, 3, ... 35.

The solution plan (algorithm) is pretty similar to the one that solved the previous problem. The differences are a sample size of 40,000 at each point in the simulation (instead of 50,000 before) and increasing the .025 width to .03 for the mode=3 comparisons that are ongoing throughout the simulation.

Therefore, in the beginning of the simulation 40,000 feasible solutions are drawn 16 times starting with the whole feasible solution space in the first of the 16 passes. However, every time a better answer (less total error) is found the simulation is re-centered about this best answer so far. Then in the second pass over the feasible solution space with 40,000 new sample answers the search region is reduced in size by a factor of 1.41 ($\sqrt{2}$) in each of the 35 dimensions of these so called 35 dimensional "rectangles". Then in the next pass the dimensions are reduced again by anther factor of 1.41 and so on for 16 times (please see Figures 1 and 2 for a partial geometric and statistical representation. Then the above mentioned 16 passes are done two more times (all with real valued arithmetic) and these three best answers are compared coordinate-wise. Then any x_i whose three values are within .03 units of each other are averaged and rounded to the nearest whole number and pinned down at that value for the rest of the simulation.

Then this whole process is repeated 24 more times and the low error (hopefully a good solution) is printed at the end of this program run.

Therefore, there are 25x3x16x40,000=48,000,000 feasible solutions that are looked at and continually averaged now leading a trail across the feasible solution space to a useful answer.

In this case the solution in the pattern turns out to be:

x_1	x_2	x_3	x_4	x_5
x_6	x_7	x_8	x_9	x_{10}
x_{11}	x_{12}	x_{13}	x_{14}	x_{15}
x_{16}	x_{17}	x_{18}	x_{19}	x_{20}
x_{21}	x_{22}	x_{23}	x_{24}	x_{25}
x_{26}	x_{27}	x_{28}	x_{29}	x_{30}
x_{31}	x_{32}	x_{33}	x_{34}	x_{35}

47	100	96	60	19
91	71	21	12	95
50	69	75	36	70
94	42	31	24	72
13	54	4	60	13
64	57	81	74	35
65	77	90	56	49

All of the equation errors printed below indicate no

0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00

error at all. This seems unrealistic given those fractional exponents in the equations and the fact that only about seven significant figures in the computer language were used (regular precision) so the continual rounding probably lost a little accuracy. The author checked the first equation with a small hand held calculator that carries many more significant figures and C_1 =500175.9 instead of 500176.1 indicating a slight error of .2 in

equation 1. This is probably the case with the other 34 equations also. However, there is still a lot of statistical evidence from the simulation that this solution has a good chance of being the true optimal (least error) as long as all whole number x_i's are required for the answer.

Switching to extended precision in the simulation (12 to 15 significant figures) would require more computer time and produce more accuracy in the equation errors, but it is doubtful that it would produce a different, more accurate, answer than this one.

Therefore, to provide a little more evidence for this being the true optimal, a program was written using extended precision calculations to check the 35 equation errors. They were all small. The total error of all equations was 3.785096818 and dividing by 35 yields an average equation error of 0.108145623.

(Conley 2018) presents additional information on modal averaging using statistical optimization with a mode of two instead of the mode of three used here.

(Wong 1990) and (Wong 1996) present multi stage Monte Carlo optimization (MSMCO) or statistical optimization applications using fewer variables and no averaging.

DISCUSSION

Even though we have powerful computers almost everywhere in the 21st century, statistical optimization or multi stage Monte Carlo optimization (MSMCO) is very computer intensive. Therefore, any steps that can be taken to reduce the amount of simulation run time is welcome as long as it does not compromise the chances of finding a useful approximate solution and/or the true optimal.

Looking at the 48 points in Figure 3, if it is desired to find a shortest route connecting all 48 points in a closed loop (a shortest route or so-called travelling salesman problem) statistical optimization works well after transforming it to a function of 48 variables. Then, if modal averaging is used the corner point sub-distances get pinned down early in the simulation because there are so few possibilities for short sub-routes there. However, in the middle of the data set (the big circle in the middle) there are so many short looking sub-routes that that part of the sub-routes in the true optimal solution may take longer to find. Therefore, maybe start the simulation with a low modal average (such as 3 or 4) and then increase the modal average to 5 or 6 when trying to figure out the difficult center sub-routes.

Thinking of the 46 capitals of European countries, Reykjavik, Lisbon, Athens and Moscow's sub-routes will get pinned down first because there are not many capitals near them, so a lower mode could be used to save time. However, Prague, Vienna and Vaduz will take longer because there are capitals in every direction from those cities (so use a higher mode). (Conley 1989) had to use modes of seven and eight (because computers were so slow then that small samples had to be drawn) on a 120 point shortest route problem connecting cities in Texas. Great circle (airplane) distances were also used.

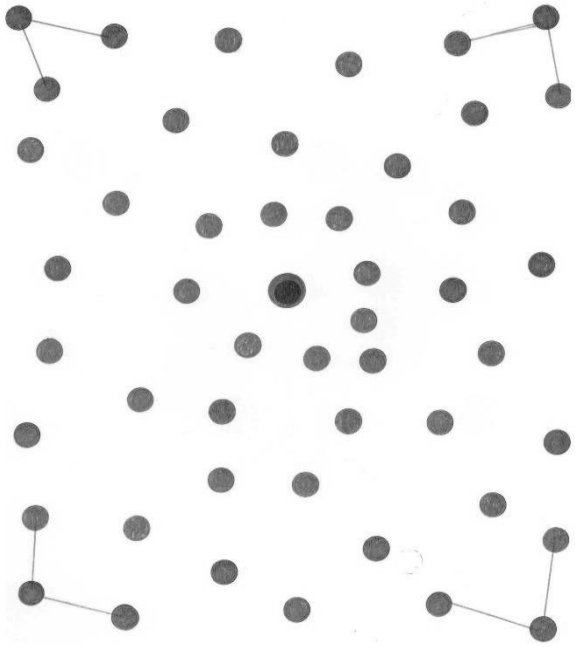


Figure 3. Note the corner points. They will show up early in the modal averaging. The routes near the big circle in the middle will be the most difficult to get right.

RENÉ DESCARTES

René Descartes, the French philosopher, scientist and mathematician in the 1600's, thought that two real number lines could be put perpendicular to each other and it would be possible to draw pictures of two variable equations. He then put a third real number line perpendicular to the x, y plane and called that variable **z** for drawing pictures of three variable equations (Descartes, 1637) and (Smith and Latham 1954).

He is called the founder of coordinate and analytic geometry and this Cartesian coordinate system that math,

science and engineering students use is named in his honor. At that time (the seventeenth century) many mathematicians and scientists thought this would lead to all applied mathematics problems being solvable. Four hundred years later the hypothetical question is, can this 21st century prove them correct?

Descartes's Cartesian coordinate system paved the way for the theoretical breakthroughs in multivariate linear algebra, linear programming and even least squares forecasting which produce "Linear Normal Equations" that then were solvable.

One big area left is nonlinear optimization. The barrier appears to be "local optimals" However, computers are so omnipresent in the 21st century and powerful (10^{17} calculations per second) (Smith and Loehrke 2018) that local optimals can be viewed now as road signs (not barriers) in the direction of the optimal region of the feasible solution space. So, if your 1000 simulations become stuck in local optimals, just average them all coordinate-wise and have another cup of coffee or tea while your computer closes in on the optimal or a useful approximation.

CONCLUSION

Two hypothetical nonlinear optimization problems were solved using multi stage Monte Carlo optimization (MSMCO) or statistical optimization with modal averaging (or Mode=3) pinning each x_i down coordinate-wise if the same value turned up in the preliminary approximate solutions.

The first problem of forty variables was solved for the exact solution (no equation error). However, the second problem had small equation errors because of the fractional exponents in the 35 equations and the fact that whole number answers were required and the computer was only using six or seven places of accuracy. It is still possible that the answer produced is the true optimal. Switching to extended precision arithmetic in the simulation (carrying 12-15 significant figures) might shed more light on this question.

Statistical optimization with averaging simulations main hazards on large scale nonlinear optimization problems used to be slow computers and local optimals in the 20th century. Now, in the 21st century, computers are so fast and ubiquitous, local optimals may be viewed as less of a barrier and more of a help in directing the way to the optimal solution in science, engineering and business.

REFERENCES

Anderson, D. R., Sweeney, D. J., Williams, T. A. 1999. *Statistics for Business and Economics*. 7th Edition. Southwestern College Publishing, Cincinnati, Ohio.

Anderson, T. W. 2003. *Multivariate Statistical Analysis*. 3rd Edition. Wiley and Sons, New York.

Black, K. 2014. *Business Statistics for Contemporary Decision Making*. 8th Edition. John Wiley and Sons, New York.

Conley, W. C. 1984. *Computer Optimization Techniques*. Revised Edition. Petrocelli Books, New York and Princeton.

Conley, W. C. 1989. "Flying the Skies of Texas." *Proceedings of the 1989 Summer Computer Simulation Conference*, Austin Texas, SCS, San Diego, CA, pp. 926-931.

Conley, W. C. 2018. "Statistical Optimization Applied to a Large Nonlinear System of Equations." In *Proceedings of the 2018 European Simulation and Modelling Conference.* ETI 2018 Ghent, Belgium, EUROSIS, Ostend, Belgium, pp. 11-16.

Descartes, René. 1637. *LaGeometries, en Francais.* Project Gutenberg. Ebook #26400, 2008. Editor, A. Hermann.

Smith, D. and Latham, M. 1954. *The Geometry of René Descartes.* Dover Publications, New York.

Smith, M. and Loehrke, J 2018. "200,000 TRILLION CALC-
ULATIONS PER SECOND." *USA today*, June 22, 2018, Weekend Edition, p. 1

Wong, J. Y. 1990 Computational experience with a general
programming algorithm. *Computers in Education Journal, 10, 19-28.*

Wong, J. Y. 1996. A note on optimization in integers. Inter-national Journal of Mathematical Education in Science and Technology*, 27 865-874.*

BIOGRAPHY

WILLIAM CONLEY received a B.A. in mathematics (with honors) from Albion College in 1970, an M.A. in theoretical mathematics from Western Michigan University in 1971, a M.Sc. in statistics in 1973 and a Ph.D. in mathematics-computer statistics from the University of Windsor in 1976. He has taught mathematics, statistics, and computer programming in universities for over 30 years. He is currently a professor emeritus of Business Administration and Statistics at the University of Wisconsin at Green Bay. The developer of multi stage Monte Carlo optimization and the CTSP multivariate correlation statistics, he is the author of five books and more than 230 publications world-wide. He is a senior member of the Society for Computer Simulation. Career highlights include presenting a paper at a National Aeronautics and Space Administration (NASA) conference in Houston, Texas and a speech at the NASA Goddard Space Flight Center outside of Washington, D.C. on the occasion of the tenth anniversary of their data center, Additionally, he gave a speech and

presentation on his research at the 7th International Conference on Genetic Algorithms in July 1997 in the United States.

PARALLEL SIMULATION

A CONCURRENT MULTI-TIER PRIORITY QUEUE FOR MULTITHREADED OPTIMISTIC PARALLEL DISCRETE EVENT SIMULATION

Matthew DePero
CSE Department, Miami University
Oxford, OH 45056, USA.
email: deperomm@miamiOH.edu

Dhananjai M. Rao
CSE Department, Miami University
Oxford, OH 45056, USA.
email: raodm@miamiOH.edu

KEYWORDS
PDES, Time Warp, Multithreading, Lockfree, Concurrent Pending Event Set, Threadsafe Datastructure

ABSTRACT

Enabling efficient multithreaded Parallel Discrete Event Simulation (PDES) is challenging and is contingent on the data structures used for managing the Pending Events Set (PES). Accordingly, we propose a lightweight, thread-safe priority queue called `3tSkipMT` for managing PES. Our design takes advantage of contemporary synchronization primitives, including atomics and lock-free instructions to address several critical design aspects and to ensure high performance. `3tSkipMT` has been incorporated into a significantly redesigned version of a parallel simulator called MUSE. The effectiveness of our proposed solution has been assessed using the standard PHOLD benchmark. The experiments show that our approach outperforms other state-of-the-art methods by achieving good, near linear speedups of ~7.5× on 8 cores.

INTRODUCTION

Rapid advancement in hardware has catalyzed the widespread use of optimistic Parallel Discrete Event Simulation (PDES). Optimistic PDES have grown in importance because they provide improved performance over conservative simulations in many scenarios Jafer et al. (2013). Moreover, the marked increase in the number of CPU-cores in machines, some with as many as 72 cores, has spurred research into enabling efficient, multithreaded optimistic PDES by leveraging shared-memory parallelism. Shared-memory parallelism offers two key advantages over conventional distributed-memory parallelism. First, it eliminates communication overheads associated with message-passing in PDES. Figure 1 illustrates the observed overhead of message-passing in a PDES. Even with contemporary Cross Memory Attach (CMA) capabilities and effective partitioning, with increasing interaction between processes, the communication overheads exceed 65% as shown in Figure 1. This conspicuous overhead can be significantly reduced through shared-memory PDES realized via multithreading (Rao, 2018). Second, multithreading reduces memory requirements by allowing a single copy of an event to be readily shared between multiple threads.

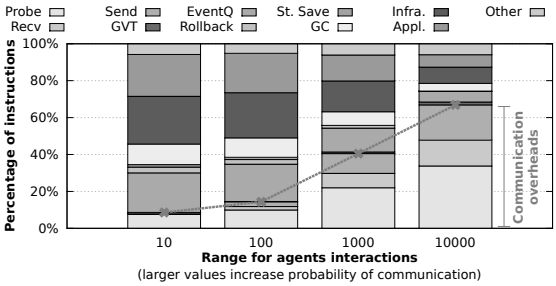


Figure 1: Communication overheads in PDES

Motivation: Challenges with multithreaded PDES

In practice, realizing the aforementioned two advantages of multithreaded PDES is challenging due to a myriad of contention, consistency, and performance issues with multithreading. A crucial issue is efficiently managing the Pending Event Set (PES), *i.e.,* set of timestamped events to be processed chronologically. The PES is the central data structure that experiences rapid enqueue, dequeue, and cancellation operations concurrently from many threads. Avoiding race conditions leads to heavy contention, thereby rapidly diminishing the realized speedup (Ianni et al., 2018). Consequently, designing efficient multithreaded PES data structures for PDES is a challenge and is an active area of research (Rao and Higiro, 2019).

Contributions: Overview of proposed research

In this study, we propose and explore a novel three-tier data structure called `3tSkipMT` to enable efficient multithreaded PDES. The key advantages of its design include: ❶ multi-tiered concurrent skip lists for reducing contention within the data structure and ❷ the use of lock-free operations to minimize the computational overheads due to contention. The `3tSkipMT` has been incorporated into a significantly redesigned version of a parallel simulator called MUSE, to enable multithreaded PDES on shared-memory platforms. Our investigations highlight several critical design issues in enabling efficient PDES. The effectiveness of the proposed solution has been assessed using the well-established PHOLD benchmark. Our solution achieves close to linear speedup of up to ~7.5× on 8 cores.

BACKGROUND: PES IN TIME WARP

Time Warp is the most widely used synchronization protocol for optimistic PDES. In Time Warp, the PDES is organized as a set of Logical Processes (LPs) that exchange *virtual time* stamped events. The Pending Event Set (PES) is a priority queue that manages the events. It strongly influences the overall performance of a Time Warp based optimistic PDES because it is used to perform the following four key operations – ❶ **Enqueue**: This operation adds one or more events, ❷ **Peek**: Identify the LP with the lowest timestamped event to be scheduled for processing, ❸ **Dequeue**: In contrast to peek, this operation removes processed events, and ❹ **Cancel**: This operation is used as part of rollback recovery process in Time Warp. This investigation aims to design an efficient, concurrent PES datastructure and associated algorithms.

RELATED RESEARCH

The Pending Event Set (PES) is the data structure for managing events to be processed. Hence, there is a breadth of literature offering a variety of PES implementations, each with its own pros and cons. One of the original data structures used for this purpose was the Calendar Queue (CQ) by Brown (1988). It uses buckets representing fixed span of virtual time to store events. A popular variant of CQ is the Ladder Queue (`LadderQ`) proposed by Tang et al. (2005). It outperforms several popular structures and the CQ by reducing of bucket management overheads by splitting buckets as needed. A key limitation of these queues is that they are primarily designed for sequential simulation and not PDES (Higiro et al., 2017). For multithreaded optimistic PDES, Gupta and Wilsey (2014) explore the use of lock-free queues for the bottom of the `LadderQ` used for managing pending events. The 3-tier Heap (`3tHeap`) for PES management was recently proposed by Rao and Higiro (2019). The `3tHeap` has shown to achieve good performance gains over other recent data structures, in a broad range of settings. However, the `3tHeap` is designed or a single threaded operations and does not permit concurrent operations from multiple threads. In contrast, we propose a novel `3tSkipMT` data structure that enables efficient concurrent multithreaded operations for high performance optimistic PDES.

Wang et al. (2014) discuss issues of enabling effective, multithreaded PDES and show performance improvements on a variety of CPU architectures. Multithreaded optimistic PDES have been investigated and reported by a team of researchers (Ianni et al., 2018). Their simulator is called Ultimate Share Everything (USE). The USE simulator utilizes a conflict-resilient, lock-free calendar queue. Our `3tSkipMT` is similar in that it uses a lock-free approach but in contrast, it uses a multi-tiered approach along with a concurrent skip list. Similar to this research, every one of the aforementioned investigations used the PHOLD performance benchmark for assessments.

MULTITHREADED PARALLEL SIMULATOR

In this study we have enabled multithreaded, Time Warp synchronized PDES by significantly redesigning an existing simulator called MUSE Rao (2018). We have adopted the approach of enhancing MUSE for several reasons, namely – ❶ MUSE already provides a well-designed Application Program Interface (API) for modeling, ❷ it includes the PHOLD benchmark easing verification and validation, ❸ MUSE's already provides the infrastructure for optimistic PDES, and ❹ MUSE already includes `3tHeap` for performance comparisons.
Figure 2 illustrates an overview of the redesigned multithreaded MUSE. A simulation consists of a given number of simulation-kernel threads. Each kernel-thread is responsible for enabling a Logical Process (LP) to operate on its next set of events. Note that at any given time each thread operates on only one LP using an independent subset of events, thereby minimizing contention. However, each thread can concurrently enqueue future events to any LP via the centralized (*i.e.,* manages events for all LPs) `3tSkipMT`. The `3tSkipMT` enables multiple threads to concurrently operate on independent but varying subset of LPs. This scheme provides an inherent load balancing scheme as opposed to fixed assignment of LPs to threads. Moreover, it reduces the inherent time differences between LPs, thereby decreasing the probability of causal violations. Nevertheless, concurrent processing of events can cause rollbacks during PDES. The cancel operation on the `3tSkipMT` enables efficiently performing the necessary rollback operations.

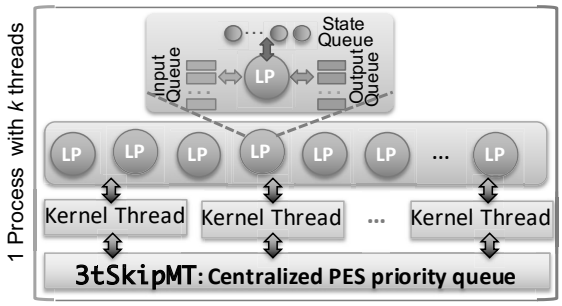


Figure 2: Overview of multithreaded MUSE PDES

Global Virtual Time (GVT) computation

In an optimistic PDES, states saved to enable rollback are garbage collected based on GVT. With our multithreaded simulations, we maintain a thread-local estimate of GVT, and the minimum of these estimates is used to compute GVT. In our proposed multithreaded PDES scheme, a single copy of an event is shared between sending and receiving LPs that could be running on different threads. Consequently, to minimize contention, we have implemented a Dual Reference Counter (DRC) approach to facilitate garbage collection. The sending and receiving LPs update two separate counters, thereby avoid race conditions. An event is deleted only when both counters are zero, indicating no thread is holding a reference to the event.

DESIGN AND IMPLEMENTATION OF `3tSkipMT`

Our proposed data structure for managing the Pending Event Set (PES) is called `3tSkipMT`. It enables concurrent and efficient operations for event management, namely enqueue, peek, dequeue, and cancel. These 4 operations have very different temporal, spatial, and concurrency characteristics. Consequently, inpired by the `3tHeap` (Rao and Higiro, 2019), we have used a 3-tiered design for `3tSkipMT` as shown in Figure 3. The first tier is a priority queue of LPs, with priority being determined by the lowest timestamp event to be processed. This tier consists of a fixed number of LPs, but their order continuously changes as events are enqueued and dequeued in lower tiers. The second tier consists of a list of buckets, ordered based on the timestamp of events in a bucket. Each bucket contains a set of concurrent events, *i.e.,* events at the same timestamp, constituting the third tier. The third tier is implemented using a standard C++ `std::vector` while the other two tiers use a concurrent skip list is discussed next.

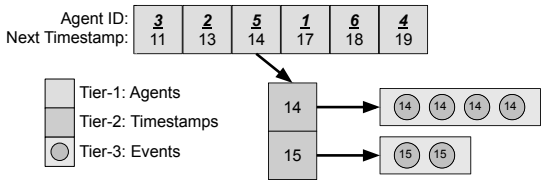


Figure 3: Overview of 3 tiers in `3tSkipMT`

In a multithreaded PDES, the first two tiers of the `3tSkipMT` will experience a significant number of concurrent operations. Consequently, we have adapted a skip list proposed by Lindén and Jonsson (2013) as it has shown to have good performance under very high contention by making balancing decisions that are agnostic to the state of the queue. Figure 4 illustrates the overview of our Concurrent Skip List (CSL). Our CSL is designed to handle simultaneous insert and delete operations without data loss, duplication, or any inaccuracy. There are two important regions of the CSL: ❶ the upper level `next[i]` (`i` > 0) skip links, are used to reduce search complexity when traversing the queue top-to-bottom and ❷ the bottom level `next[0]` represents the true priority order of events in the queue. The "height" of a level (l) for a single node is chosen probabilistically ($p(l)$) via a geometric distribution $p(l) = 0.5 \cdot p(l-1)$. Starting with a 100% probability of reaching the bottom level, for a given level (l), the likely hood of a node reaching the level above it is reduced by half. Hence, traversing links from the top down creates an effective binary search for long lists.

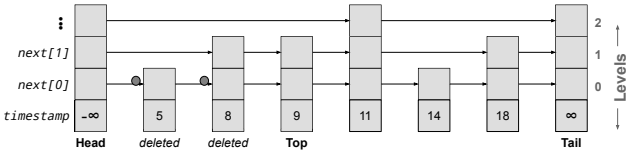


Figure 4: Concurrent Skip List (CSL) used for tiers 1 & 2

Enqueue, Peek, & Dequeue operations

As a result of the CSL design, we can concurrently insert and delete while maintaining causality by manipulating pointers on level-0. Next, the higher levels are suitably restructured to decrease search time but without impacting lower levels. These operations are performed using atomic compare-and-swap (CAS) instructions. Conflicts due to concurrent modifications cause the CAS instructions on one of the threads to fail, prompting those threads to then retry the operation.

Cancel operations: Issues & implementation

A novel modification that has been introduced in our CSL is the ability to delete an arbitrary key from the middle of the queue for event cancellation and rescheduling during rollbacks. This feature has been accomplished by introducing a second atomic operation during event deletion – *i.e.,* in addition to flipping a deleted-bit (to detect concurrent operations) on the `next[]` pointers, we also clear the event pointer (thereby logically deleting the node) contained in the node as part of the atomic CAS instruction. This enhancement allows threads to *logically* delete an element from the middle of the queue. However, these nodes are not immediately deleted but are allowed to naturally propagate to the top of the CPQ where they are detected and deleted. This design is a tradeoff between slight increase in memory footprint versus increased concurrency.

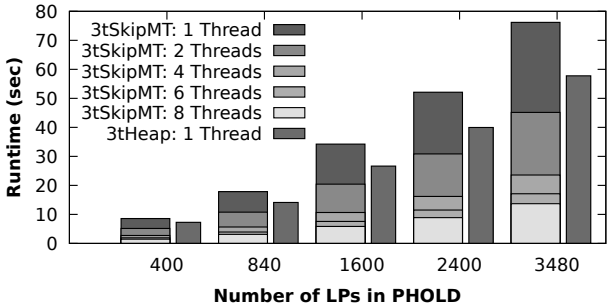
EXPERIMENTS & RESULTS

Experimental assessment of our multithreaded parallel simulator have been conducted a well-established synthetic benchmark called PHOLD. It is widely used in PDES investigations because it has shown to effectively emulate the steady-state phase of a broad range of real world models (Rao and Higiro, 2019). A key advantage of PHOLD is that it includes a variety of settings that can be used to configure the benchmark to mimic characteristics of different real world models. From our prior research (Rao and Higiro, 2019), we have used generalized sensitivity analysis and we have identified that two PHOLD parameters play an influential role this research, namely: ❶ the number of LPs which influences the net number of events in the simulation and ❷ each event's granularity, *i.e.,* CPU-time needed to process an event. Accordingly, we explore these influential parameters in the following subsection for performance assessments. In addition, PHOLD enables comparison with another Time Warp based simulator called Ultimate Share-Everything (USE) as discussed in the final subsection.

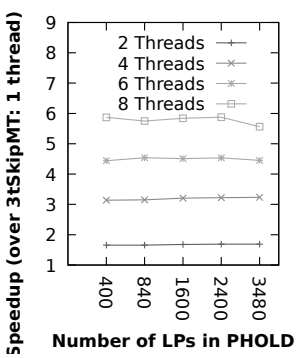
Scalability & performance results

The chart in Figure 5(a) illustrates the observed average runtimes with a different number of threads and different number of Logical Processes (LPs) in the PHOLD benchmark. The chart also plots the average runtimes using the `3tHeap`, which is used as the reference for comparisons in this specific analysis. We have used `3tHeap` for comparison as it

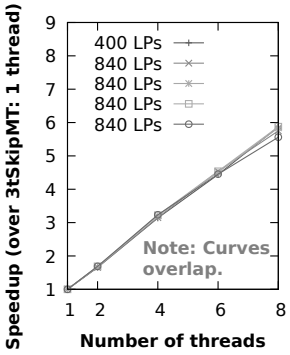
is currently one of the fastest datastructures for sequential PDES (Rao and Higiro, 2019). As illustrated by the chart in Figure 5(a), the performance of `3tSkipMT` is lower than the `3tHeap` by about 15% to 24%. This difference reflects two specific overheads of the `3tSkipMT`, namely: ❶ the overhead of using lock-free instructions which are known to degrade Instructions Per Clock-cycle (IPC) of the CPU (Rao, 2018), and ❷ slight increase in cache misses due to the use of linked lists in the `3tSkipMT`. However, as the number of threads is increased, the computation is spread across corresponding number of CPU-cores, thereby decreasing the net runtime of the simulations.



(a) Simulation runtime comparisons



(b) LPs vs. speedup



(c) Threads vs. speedup

Figure 5: Performance & scalability of `3tSkipMT`

The plots in Figure 5(b) and Figure 5(c) shows very good scalability – *i.e.,* as the number of threads are increased the performance of the simulation proportionally increases. Such scalability profile reflects the two key advantages of the design of `3tSkipMT` – *i.e.,* ① reduced contention within the data structure and ② the use of lock-free operations to minimize the computational overheads due to contention. However, the observed speedup is slightly below the theoretical speedup because of thread contention on shared memory regions.

Influence of event granularity

Event granularity (*i.e.,* average CPU-time for processing an event) is an important factor that indirectly influences the overall multithreaded performance as it determines contention for shared memory within the `3tSkipMT`. Higher granularity values reflect scenarios where threads are operating independently for longer duration. Figure 6 plots the impact of granularity on multithreaded simulation runtimes. As illustrated by the curves in Figure 6, the `3tSkipMT` provides good multithreading performance as the event granularity increases. The bottlenecks due to increased contention are prominent only at zero granularity, where multithreading is slower than the single-threaded simulation. However, zero granularity scenario is purely a theoretical (used for analysis) and in practice event angularities are well above zero. These experiments clearly show that the design of `3tSkipMT` is effective at minimizing contention and improving concurrency.

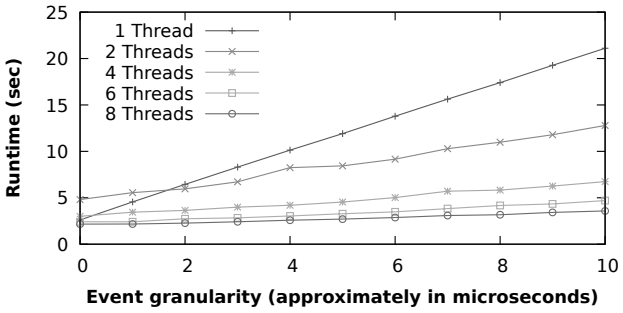


Figure 6: Influence of per-event granularity

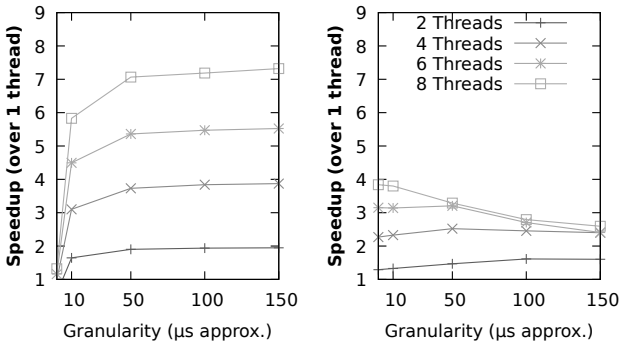
Comparative scalability analysis

Scalability is a measure of how well a simulator is able to utilize added resources, *i.e.,* CPU cores. In order to determine the scalability of `3tSkipMT` in a larger context, we have compared its scalability to another simulator called Ultimate Share-Everything (USE) by Ianni et al. (2018). It uses a Conflict Resilient Calendar Queue (CRCQ) that is lock-free to enable multithreaded, Time Warp synchronized PDES. However, it must be noted that USE is a completely different simulator when compared to our experimental testbed. Consequently, a direct comparison of the runtimes of two systems is not meaningful. Instead, we only compare the relative behaviors of the two systems from the perspective of scalability.

The charts in Figure 7 illustrate a comparison of the scalability of `3tSkipMT` versus CRCQ/USE for different granularity settings. In these experiments, `3tSkipMT` continues to exhibit good scalability as the number of threads are increased. For example, with 8 threads, `3tSkipMT` provides slightly over 7× speedup, at granularity > 50 μs. The scalability profile of `3tSkipMT` is substantially better when compared to the scalability profile observed for CRCQ in Figure 7(b). The improved scalability profile is of `3tSkipMT` highlights one of the key benefits this novel data structure for multithreaded PDES.

CONCLUSIONS & FUTURE WORK

Shared-memory parallelism realized via Multithreading is the primary approach used for conducting Parallel Discrete Event Simulation (PDES) on contemporary CPUs. In prac-



(a) 3tSkipMT (b) CRCQ/USE

Figure 7: Influence of per-event granularity

tice, realizing good performance from multithreaded PDES is challenging due to a myriad of contention, consistency, and performance issues with multi-threading. This paper presented our efforts to enable efficient, scalable optimistic PDES on shared-memory platforms. The paper presented the design, design rationales, and implementation of our multithreaded, Time Warp synchronized, general purpose parallel simulator called MUSE. The cornerstone of our multithreaded approach is a concurrent priority queue called 3tSkipMT that is used to manage the Pending Event Set (PES). It is thread-safe – that is, multiple threads can concurrently operate on it without causing race conditions. Its design permits a free thread to immediately schedule an LP while providing an inherent load balancing scheme and decreasing the probability of causal violations. These proprerties have been realized using a 3-tier design, with first two tiers have been implemented using a novel Concurrent Skip List (CSL). The CSL operations have been implemented using contemporary lock-free operations to minimize contention overheads. In addition, we have developed a novel method for logically deleting arbitrary keys from the middle of the CSL.

We have extensively verified and validated our multithreaded parallel simulator using the PHOLD benchmark. The experiments show that our methodology yields almost linear scalability characteristic. We observed a very good ~7.5× runtime reduction using 8 CPU-cores, particularly for event granularity > 50 μs. Our experiments highlight the effectiveness of our methodology for enabling efficient multithreaded optimistic PDES.

Future work

We are continuing our investigations to further improve the performance related aspects – ❶ minimize pointers to improve CPU-cache performance; ❷ explore lock-free algorithms for faster restructuring the CSL; ❸ enable distributed, multithreaded PDES on supercomputing clusters. The aforementioned enhancements would enable realizing efficient PDES on the next generation of exascale supercomputers that are already on the horizon.

REFERENCES

Brown R., 1988. *Calendar Queues: A Fast 0(1) Priority Queue Implementation for the Simulation Event Set Problem*. *Commun ACM*, 31, no. 10, 1220–1227. ISSN 0001-0782. doi:10.1145/63039.63045.

Gupta S. and Wilsey P.A., 2014. *Lock-free Pending Event Set Management in Time Warp*. In *Proceedings of the ACM SIGSIM PADS*. ACM, New York, NY, USA. ISBN 978-1-4503-2794-7, 15–26.

Higiro J.; Gebre M.; and Rao D.M., 2017. *Multi-tier Priority Queues and 2-tier Ladder Queue for Managing Pending Events in Sequential and Optimistic Parallel Simulations*. In *Proceedings of the 2017 ACM SIGSIM Conference on Principles of Advanced Discrete Simulation*. ACM, ACM, New York, NY, USA, SIGSIM-PADS '17. ISBN 978-1-4503-4489-0, 3–14. doi:10.1145/3064911.3064921.

Ianni M.; Marotta R.; Cingolani D.; Pellegrini A.; and Quaglia F., 2018. *The Ultimate Share-Everything PDES System*. In *Proceedings of the 2018 ACM SIGSIM Conference on Principles of Advanced Discrete Simulation*. ACM, New York, NY, USA, SIGSIM-PADS'18. ISBN 978-1-4503-5092-1, 73–84. doi:10.1145/3200921.3200931.

Jafer S.; Liu Q.; and Wainer G., 2013. *Synchronization methods in parallel and distributed discrete-event simulation*. *Simulation Modelling Practice and Theory*, 30, 54–73. ISSN 1569-190X.

Lindén J. and Jonsson B., 2013. *A Skiplist-Based Concurrent Priority Queue with Minimal Memory Contention*. In R. Baldoni; N. Nisse; and M. van Steen (Eds.), *Principles of Distributed Systems*. Springer International Publishing, Cham. ISBN 978-3-319-03850-6, 206–220.

Rao D.M., 2018. *Performance Comparison of Cross Memory Attach Capable MPI vs. Multithreaded Optimistic Parallel Simulations*. In *Proceedings of the 2018 ACM SIGSIM Conference on Principles of Advanced Discrete Simulation*. ACM, New York, NY, USA, SIGSIM-PADS '18. ISBN 978-1-4503-5092-1, 37–48. doi:10.1145/3200921.3200935.

Rao D.M. and Higiro J.D., 2019. *Managing Pending Events in Sequential and Parallel Simulations Using Three-tier Heap and Two-tier Ladder Queue*. *ACM Trans Model Comput Simul*, 29, no. 2, 9:1–9:28. ISSN 1049-3301. doi:10.1145/3265750.

Tang W.T.; Goh R.S.M.; and Thng I.L.J., 2005. *Ladder Queue: An O(1) Priority Queue Structure for Large-scale Discrete Event Simulation*. *ACM Trans Model Comput Simul*, 15, no. 3, 175–204. ISSN 1049-3301.

Wang J.; Jagtap D.; Abu-Ghazaleh N.; and Ponomarev D., 2014. *Parallel Discrete Event Simulation for Multi-Core Systems: Analysis and Optimization*. *IEEE Transactions on Parallel and Distributed Systems*, 25, no. 6, 1574–1584. ISSN 1045-9219. doi:10.1109/TPDS.2013.193.

PARALLEL VERIFICATION METHODOLOGY OF RECONFIGURABLE HIERARCHICAL TIMED AUTOMATA

Roufaida Bettira
National Institute of Applied Sciences and Technology,
University of Carthage, Tunis, Tunisia
email: roufaida.bettira@gmail.com

Laid Kahloul
LINFI Laboratory, Computer Science Department,
University of Mohamed Khider, Biskra, Algeria
email: laid.k.b@gmail.com

Mohamed Khalgui
National Institute of Applied Sciences and Technology,
University of Carthage, Tunis, Tunisia
email: khalgui.mohamed@gmail.com

KEYWORDS
Discrete event control system, Verification, Timed automata, Reconfiguration, Hierarchy, Parallelism.

ABSTRACT

This paper deals with formal verification of hierarchical reconfigurable discrete event control systems (RDECS) using reconfigurable hierarchical timed automata (RHTA) which is an extension of timed automata. These large systems require an improved verification approach where parallelism is introduced. We propose in this paper, a parallel methodology for RHTA formal verification with reduced time and memory space. The contribution is applied to a real case study (smart grids) since they are characterized with reconfigurability and hierarchy features to show the gains.

INTRODUCTION

Failures in critical systems such as medical or aerospace systems may lead to fatal consequences. Therefore, it is necessary to verify if some requirements are ensured before validation. Formal verification is a technique based on mathematical tools that checks the designed system. Reconfigurable hierarchical systems have hierarchical changeable structures and behaviors over time. Reconfiguration is required either to response requirements of users or to handle unexpected hardware failures. Hierarchical modeling is useful since it is truly applicable to real-life embedded systems, it specifies large systems in different levels and it allows the reuse of components.
Timed automata (TA) represent a formalism for modeling and verification of real-time systems. Verification of properties is done using model-checker which receives two inputs, the system model and the set of properties to be checked specified in a temporal logic. To introduce reconfigurability and hierarchy to TA models, the extension reconfigurable hierarchical timed automata (RHTA) is proposed in (Roufaida et al. 2019). RHTA is a structure of hierarchical behavior graphs where each graph represents a configuration and a node is represented by one basic component (a TA model) or TA network. A behavior graph is reconfigured to another one by applying reconfiguration functions.
In the literature, several papers tend toward the verification of hierarchical systems. In (Alur and Yannakakis 2001), authors propose new techniques and algorithms for verification of hierarchical state machines (HSMs). The work in (Sklyarov and Skliarova 2008) presents a novel model and method for the design of parallel hierarchical finite state machines (PHFSM) however, there is no contribution for the verification of such models. In (La Torre et al. 2008), authors aim for studying an extension of an HSM called scope-dependent hierarchical state machines (SHSMs). They prove that verification can be more efficient since SHSM models are more succinct than HSMs. DIVINE is a tool for LTL model checking and reachability analysis of discrete distributed systems. In (Barnat et al. 2010), a new release of DIVINE is introduced such that parallel features are considered. The work in (Lin and Li 2014) handle the state space explosion problem in verification of complex analog circuits by proposing a parallel hierarchical reachability analysis method.
An Improved verification of RHTA is also proposed in (Roufaida et al. 2019) where redundancies between different configurations are considered. However, the verification process inside the configuration itself is executed sequentially. This paper proposes a parallel approach for RHTA verification. The idea is that verification of basic components and generation of their interfaces are distributed over set of processors and are effectuated in parallel. After that, initial behavior graph is divided in depth and obtained branches are also verified in parallel. For other graphs (configurations), similarities are always considered. Parallel verification proves gains in terms of required time compared to the sequential one.
The remainder of this paper is organized as follows. The second section presents the background concerning the

extension RHTA and its sequential verification methodology. The third section provides a new methodology of RHTA verification including parallel strategy. The fourth section handles smart grids as a real case study to apply and evaluate the proposed contribution. Finally, the fifth section concludes this paper.

PRELIMINARIES

The proposed parallel verification methodology is based on the extension RHTA proposed in (Roufaida et al. 2019). Modeling and verification of an RHTA are briefly recalled in this section.

Reconfigurable Hierarchical Timed Automata

RHTA (Roufaida et al. 2019) are an extension of timed automata (Bengtsson and Yi 2003, Bouyer et al. 2008). An RHTA is a structure RA= (BC, BG, R), where (i) BC is a finite set of basic components, (ii) $BG = \{g_0, g_1, ..., g_n\}$ is a set of n hierarchical behavior graphs representing the different behaviors (configurations) performed by the modeled system, and (iii) R is a finite set of m reconfiguration functions $R = \{r_1, r_2, ..., r_m\}$. A reconfiguration function r is a structural modification that transforms one configuration to another after the fulfillment of certain preconditions.
A basic component in set BC is represented as a timed automaton. A hierarchical behavior graph is a tuple g= (V, V_0, T, Lab, $Prob$, F) where,

- V: is a finite set of vertices (nodes) such that, $v \in V$ is represented by one or several basic components.
- $V_0 \subset V$: is a set of initial vertices.
- $T : V \to V$ is the decomposition transition relation between two vertices from two successive levels.
- Lab: $2^V \to \{AND, OR\}$ is a labelling function mapping a subset of vertices outgoing from the same vertex to one of the labels "AND" or "OR".
- $Prob$: $T \to]0, 1]$ is a probabilistic transition function that maps each transition to a real number in interval $]0, 1]$, where 1 is the default probabilistic value for a transition.
- $F \subset V$: is a finite set of end vertices in the last level of the hierarchy.

A behavior graph represents one configuration of the hierarchical system (as seen in Fig.1):
Formally, a reconfiguration function r is a couple ($Cond$, m) where,

- $Cond \in \{true, false\}$ is the precondition of **r**.
- $m : (g, v) \to (g', v')$ is the structure modification instruction that transforms one graph (i.e., configuration) g to another g', where $g, g' \in BG$.

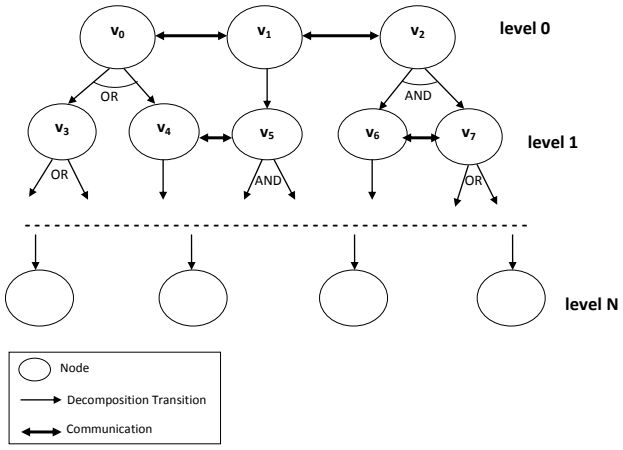


Figure 1: A configuration of RHTA.

The set of fundamental structure modification instructions of RHTA is detailed in Table 1 such that $v_1, v_2 \in V$ are two vertices, $M \in BC$ is a basic component, $x \in]0, 1]$ is a real number, and $X \subseteq V$ is a subset of vertices. The concatenation between basic modification instructions is denoted by $\dot{+}$ to build complex modification instructions.

Table 1: Fundamental Structure Modification Instructions of RHTA.

Instruction	Symbol
Add M to vertex v_1	new(M, v_1)
Remove M from vertex v_1	rmv(M, v_1)
Add vertex v_1	new(v_1)
Remove vertex v_1	rmv(v_1)
Insert a transition from v_1 to v_2 with the probability value x	new(v_1, x, v_2)
Remove transition from v_1 to v_2	rmv(v_1, v_2)
Modify the probability value of the transition from v_1 to v_2 by x	mdf(v_1, x, v_2)
map the label AND to X	$AND(X)$
map the label OR to X	$OR(X)$

Sequential Verification Methodology of RHTA

Definition 1: The semantics of TA is given through a transition system (denoted by TS, it is also named timed transition system) (Bengtsson and Yi 2003, Bouyer et al. 2008) where a state is a pair $\langle l, u \rangle$, l is the current location and u is a function providing the current values of clocks ($u : C \to R_{\geq 0}$). There are two types of transitions between states. The automaton may either delay for some time (i.e., a delay transition), or follows an enabled edge (i.e., an action transition).

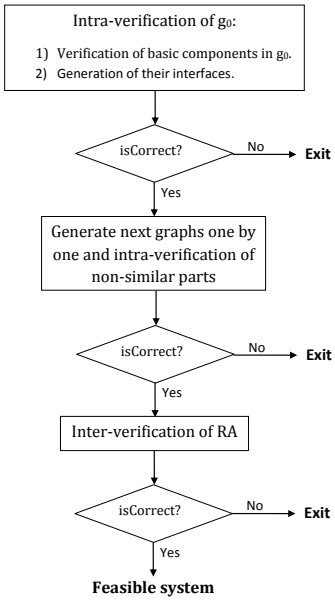


Figure 2: Organizational Chart of RHTA sequential Verification.

Definition 2: Semantics of hierarchical behavior graphs $g \in BG$ in RHTA is given through hierarchical transition systems (HTS). An HTS is defined by HTS= (S, Tr) as reported in (Roufaida et al. 2019) where, S is a set of states representing the different TSs of automata in each $v \in V$ and $Tr \subseteq (S \times S)$ is the decomposition transition function between those states such that $Tr \subseteq T$.
An interface of a TA model is defined as part of this model where only input and output actions are considered. The verification methodology idea is about considering similarities between different configurations, redundant parts are eliminated in each generated HTS. Basic components are verified first and their interfaces are generated, then initial configuration, after that verification of other generated configurations, and finally the entire system. We show in Fig. 2 different steps followed for the sequential verification of the structure RA= (BC, BG, R) as reported in (Roufaida et al. 2019).

PARALLEL VERIFICATION METHODOLOGY OF RHTA

Motivation

In our previous work (Roufaida et al. 2019), we have proposed the extension RHTA for formal modeling and verification of RDECSs. The proposed verification methodology is sequential. In this paper, we propose a parallel methodology for RHTA verification and we define its complexity in terms of execution time. Parallel verification is useful since RHTA deals with large amounts of data. We note that computational tree logic (CTL) (Boucheneb et al. 2009) is used to specify several functional properties such as correctness, deadlock-freeness, reachability, and safety.

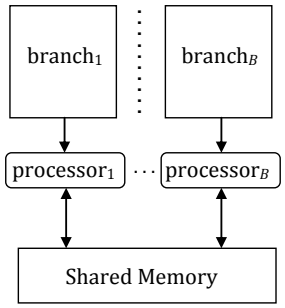


Figure 3: Parallel Architecture for Verification of a Configuration in RHTA.

Formalization

Definition 3: We call a branch in a $g \in BG$ each chain of nodes starts with an initial vertex $v \in V_0$, finishes with an end vertex $v \in F$, and includes a vertex from each level related with decomposition transitions.
Parallel verification methodology of RHTA consists of divide our hierarchical model in depth into set of branches where B is the number of these branches. Fig. 3 presents the proposed parallel architecture for the verification of one hierarchical behavior graph (configuration) in RHTA. Indeed, only initial configuration is verified, redundancies in other configurations are eliminated in the verification task. The architecture is constituted of B processors in worst case that execute the B branches in parallel. Processors are connected to a common memory to guarantee execution precedence between the processes during parallel verification.

Verification of basic components and generation of interfaces: Before the composition of basic components and verification of each graph (configuration), the set BC of basic components is verified first to check the satisfiability of the concerned properties as soon as possible. Then, the interface of each basic component is generated.
Set BC is distributed over B processors which means that $\frac{|BC|}{B}$ components are verified and $\frac{|BC|}{B}$ interfaces are generated in parallel.

Verification of initial configuration: Let us consider *vec* as a vector to store the status of each node (vertex) in initial graph whether "verified" or "not verified". Initially, all nodes are not verified. The proposed parallel verification of initial configuration is carried out following this strategy:

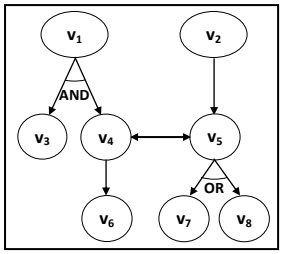


Figure 4: An example of a configuration designed with RHTA.

Processors \ instant	t_1	t_2	t_3
pr_1	v_1	v_3	v_8
pr_2	v_2	v_5	v_7
pr_3		v_4	v_6

Figure 5: Nodes verification scheduling.

1. all independent branches (i.e., branches which have not any common nodes) starting from left in the hierarchical graph are verified in parallel. Correspondent processors change the status of nodes in these branches to "verified".

2. in the same time other branches are verified in parallel without re-verifying common nodes that have already been verified. The status of each node is checked before the verification of this node.

Example: We present in Fig. 4 a simple example of one configuration modeled with RHTA. In this case, the graph can be divided to $B = 4$ branches which are verified following the proposed strategy as depicted in Fig. 5. We start with verification of independent branches (v_1, v_3) and (v_2, v_5, v_7) which are attributed to two processors pr_1 and pr_2, respectively. The two other branches (v_1, v_4, v_6) and (v_2, v_5, v_8) are verified in parallel without re-verifying common vertices v_1, v_2, and v_5.

Verification of other configurations: Once initial configuration is verified other ones are generated one by one by applying set of transformation rules as reported in (Roufaida et al. 2019). Only new parts are verified using same parallel strategy as initial configuration verification.

Complexity and Gains

We discuss in this subsection, the complexity of the proposed parallel verification methodology compared to sequential one.
We recall that the entire sequential verification approach of RHTA passes with these steps: 1) verification of basic components in each configuration which gives at the end, the verification of the whole set BC, $nbc = |BC|$, 2) generation of interfaces of basic components, f is the average time for calculating an interface, 3) verification of the initial behavior graph, $s = |V|$ is the number of vertices in g_0, and 4) verification of each reduced $g \in BG$, $k = |BG|$ and r is the average number of applied transformation rules on a graph to get another one. The complexity of the entire sequential verification of a RDECS modeled with RHTA can be determined in time with

$$\mathcal{O}(nbc * e^u + nbc * f + 1 * e^u * s + (k-1) * r)$$

where: 1) $\mathcal{O}(e^u)$ is the complexity of TA's verification using the classical model checker, the complexity as reported in (Bouyer et al. 2008, Ouaknine and Worrell 2008) depends on the time constraints (guards) and actions which influence the state space explosion (i.e., u is the clock valuation function). This complexity is used nbc times to verify the set of basic components which gives ($nbc * e^u$), also used just once for the verification of each vertex (TA network) in the initial behavior graph which gives ($1 * e^u * s$); 2) $\mathcal{O}(nbc * f)$ is the complexity that determines the time for generating the set of BC interfaces; and 3) $\mathcal{O}((k-1) * r)$ is the complexity for the verification of the $k-1$ other behavior graphs obtained from g_0.
The proposed parallel approach improves the required time for the verification of a hierarchical reconfigurable system. The verification cost of set BC and generation of their interfaces is reduced $\frac{nbc}{B}$ times since they become parallel operations which gives $\mathcal{O}(\frac{nbc}{B}(e^u + f))$ instead of $\mathcal{O}(nbc * e^u + nbc * f)$. In addition, different branches of the initial configuration are proposed to be checked in parallel, thereby the complexity $\mathcal{O}(e^u * s)$ is reduced to $\mathcal{O}((e^u * s)/B)$. The complexity of the entire parallel verification of a RDECS modeled with RHTA can be determined in time with

$$\mathcal{O}(\tfrac{nbc}{B}(e^u + f) + (e^u * s)/B + (k-1) * r)$$

EXPERIMENTATION

Motivation

Smart grids (SG) are systems which are widely handled in last years, they are characterized by their hierarchical structure and their reconfigurable behavior, thereby RHTA extension is fitted for their design. First, we model the smart grid hierarchy using RHTA and we apply its sequential verification methodology for the verification of this system. Then, we apply the parallel verification approach proposed in this paper and we show the gains. We vise a part of the traditional power transmission system (TPTS) of West Lake/North Urban Centres in Tunis (Tunisia) as a real case study(Jiang et al. 2014).

Smart Grid Formal Modeling

We consider the TPTS which is shown in Fig6, this grid is specified with an RHTA as SG= (BC_{grid}, BG_{grid}, R_{grid}), where:

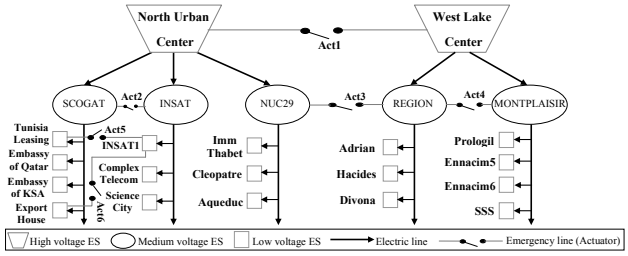


Figure 6: A part of West Lake/North Urban Center in Tunis.

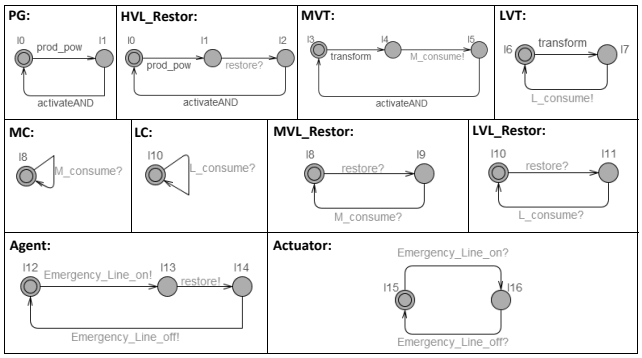


Figure 7: The set of basic components in SG.

- BC_{grid}={North Urban, West Lake, MVT1, MVT2, SCOGAT, INSAT, NUC29, REGION, MONTPLAISIR, LVT1, LVT2, LVT3, LVT4, LVT5, LC1, LC2, LC3, LC4, LC5, LC6, LC7, LC8, LC9, LC10, LC11, LC12, LC13, LC14, LC15, LC16, LC17, Agent1, Act, HVL_R, MVL_R, LVL_R } are its basic components as shown in Fig7,
- BG_{grid}={g_0, ..., g_{24}} is the set of behavior graphs (configurations) in this system, and
- R_{grid}={$r_{g_0-g_1}$, $r_{g_0-g_2}$, $r_{g_0-g_3}$, $r_{g_1-g_0}$, $r_{g_2-g_0}$, $r_{g_3-g_0}$} is set of reconfiguration functions.

This is an example of a SG where, North Urban and West Lake are two central power generations which are declared from the designed TA model (template) **PG**. MVT1 and MVT2 are two medium voltage transformers declared from **MVT**. LVT1, LVT2, LVT3, LVT4, and LVT5 are five low voltage transformers declared from **LVT**. SCOGAT, INSAT, NUC29, REGION, and MONTPLAISIR are medium consumers declared from **MC**. LC1, LC2, LC3, LC4, LC5, LC6, LC7, LC8, LC9, LC10, LC11, LC12, LC13, LC14, LC15, LC16, and LC17 are low consumers declared from **LC**. Agent1 is a declared agent from the designed TA model **Agent**, it is responsible of restoring the electrical faults by closing the emergency line (actuator) between the faulty substation and the supplier one. Act is an actuator declared from **Actuator**. HVL_R, MVL_R, and LVL_R

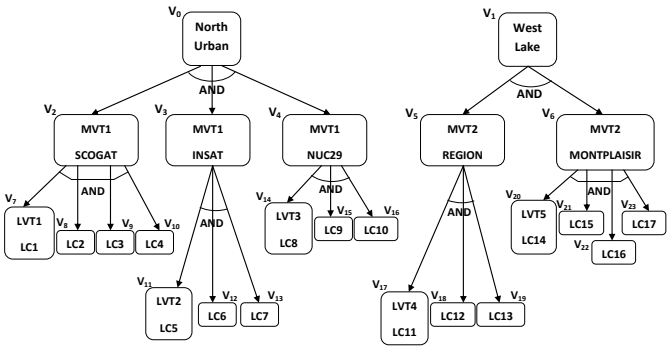


Figure 8: Initial configuration g_0.

are restored high, medium, and low voltage lines declared from the templates: HVL_Restor, MVL_Restor, and LVL_Restor, respectively.

In this study, we deal with three configuration categories beside the initial configuration g_0 such that, g_1, g_2, and g_3 are three model configurations representing the first, second, and third category, respectively. The first (resp. second and third) category are constituted of two (resp. five and seventeen) configurations depending to the number of high (resp. medium and low) voltage lines in this TPTS. The initial configuration g_0 as depicted in Fig. 8 represents the system in its normal functionality (no electrical faults), the TPTS is switched to g_1, g_2, or g_3 when a fault is occurred in one of its high, medium, or low input lines, respectively. The fallen down substation will be supplied by another electrical substation (fault restoration through actuator closure). To illustrate the three possible faults in this system, we treat the fault occurrence in North Urban Centre, REGION, and INSAT1 as examples from each level. We describe in table 2, the six considered reconfiguration functions in SG.

Smart Grid Formal Verification

We consider $\phi = AFG$ not deadlock as the deadlock-free property to be checked. The property ϕ is specified in CTL.

Sequential Verification

We apply the sequential verification methodology of RHTA as recalled in preliminaries. First of all, basic components are checked against the property ϕ and their interfaces are generated. Then, Verification needs computation of HTSs. We start by computing HTS_0 as depicted in Fig 9 where generated TSs: TS_0, TS_2, and TS_7 are represented in Fig10, Fig11, and Fig12, respectively. In this TPTS, TSs in one level have identical behaviors. For the last level, TS_7 is resulted from the merger of the vertices v_7, v_8, v_9, and v_{10}. Finally, TA model-checker is applied to each TS in HTS_0 starting

Table 2: Reconfiguration Functions of SG.

reconfig. function	Cond	m
$r_{g_0-g_1}$	fault is occurred in **North Urban Centre**	rmv(North Urban, v_0)$\dotplus$ new(HVL_R, v_0)$\dotplus$ new(Agent1, v_0)$\dotplus$ new(Act, v_1)
$r_{g_0-g_2}$	fault is occurred in **REGION**	rmv(REGION, v_5)$\dotplus$ new(MVL_R, v_5)$\dotplus$ new(Agent1, v_5)$\dotplus$ new(Act, v_4)
$r_{g_0-g_3}$	fault is occurred in **INSAT1**	rmv(LC5, v_{11})$\dotplus$ new(LVL_R, v_{11})$\dotplus$ new(Agent1, v_{11})$\dotplus$ new(Act, v_7)
$r_{g_1-g_0}$	the faulty high voltage line is repaired	new(North Urban, v_0)$\dotplus$ rmv(HVL_R, v_0)$\dotplus$ rmv(Agent1, v_0)$\dotplus$ rmv(Act, v_1)
$r_{g_2-g_0}$	the faulty medium voltage line is repaired	new(REGION, v_5)$\dotplus$ rmv(MVL_R, v_5)$\dotplus$ rmv(Agent1, v_5)$\dotplus$ rmv(Act, v_4)
$r_{g_3-g_0}$	the faulty low voltage line is repaired	new(LC5, v_{11})$\dotplus$ rmv(LVL_R, v_{11})$\dotplus$ rmv(Agent1, v_{11})$\dotplus$ rmv(Act, v_7)

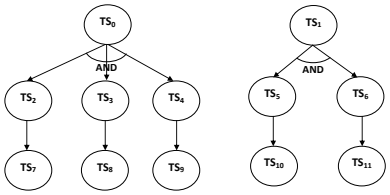


Figure 9: HTS_0.

with TS_0 to check ϕ satisfiability.
HTS_1, HTS_2, and HTS_3 are obtained after applying the reconfiguration functions $r_{g_0-g_1}$, $r_{g_0-g_2}$, and $r_{g_0-g_3}$, respectively on HTS_0. New states (TSs) in resulted HTSs are computed while the rest of states are preserved from HTS_0, thereby only TS_0, TS_5, and TS_7 are rebuilt to TS'_0, TS'_2, and TS'_7, respectively as is shown in Fig13, Fig14, and Fig15, respectively. Property ϕ is only rechecked in TS'_0, TS'_2, and TS'_7.

Parallel Verification

Application of the proposed parallel verification methodology on SG consists of dividing HTS_0 into five branches. Therefore, our parallel architecture will include 5 processors in the worst case. First, property ϕ is checked in $\frac{|BC_{grid}|}{5}$ basic components in parallel. Then, $\frac{|BC_{grid}|}{5}$ interfaces are generated in parallel. After that, we check the satisfiability of ϕ in the initial configuration branches in parallel following the scheduling

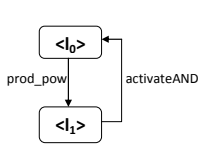


Figure 10: TS_0.

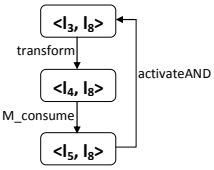


Figure 11: TS_2.

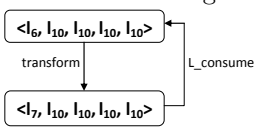


Figure 12: TS_7.

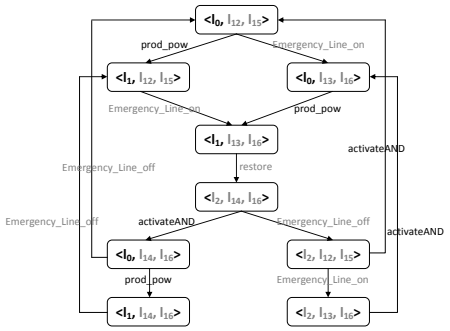


Figure 13: TS'_0.

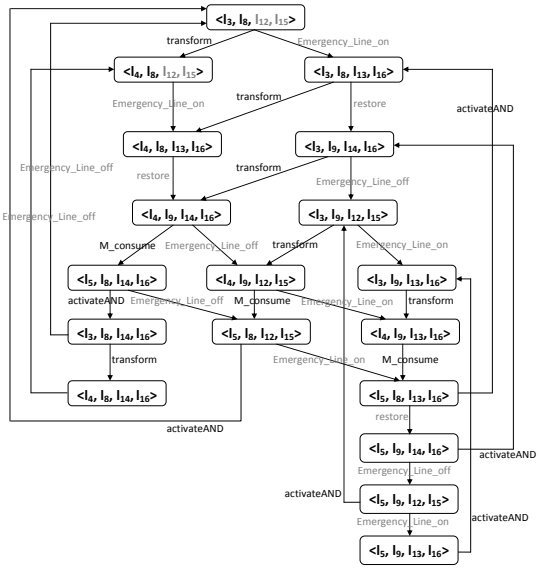


Figure 14: TS'_2.

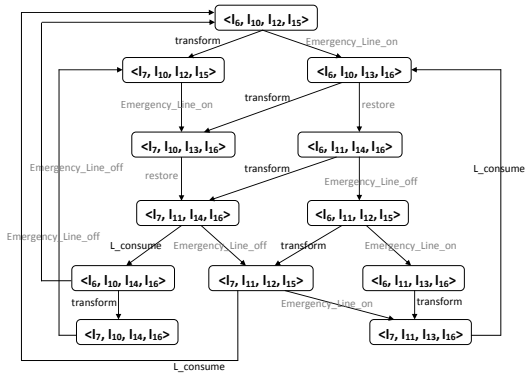


Figure 15: TS'_7.

Processors \ instant	t_1	t_2	t_3
pr_1	TS_0	TS_2	TS_7
pr_2	TS_1	TS_5	TS_{10}
pr_3		TS_3	TS_8
pr_4		TS_4	TS_9
pr_5		TS_6	TS_{11}

Independent branches

Figure 16: Verification scheduling of initial configuration.

presented in Fig. 16. Finally, HTS_1, HTS_2, and HTS_3 are generated from HTS_0 after applying the reconfiguration functions $r_{g_0-g_1}$, $r_{g_0-g_2}$, and $r_{g_0-g_3}$, respectively. property ϕ is than rechecked only in TS'_0, TS'_2, and TS'_7.

Discussion

The proposed parallel verification methodology shows important gains, especially in basic components verification and initial configuration verification stage. Improvements in generation and verification of other configurations are already proposed in our previous work (Roufaida et al. 2019) which are also adopted in this contribution. We have 36 basic components which are verified in $\frac{36}{5} \simeq 7$ time units instead of 36 time units. For HTS_0, twelve TSs are calculated in three time units instead of 12 time units using sequential verification methodology. Considering similarities, HTS_1, HTS_2, and HTS_3 are generated from HTS_0 with one single modified TS in each. One time unit is required to verify each new TS in both methods.

CONCLUSION

This paper proposes a parallel verification methodology of RHTA. Compared to the sequential method, gains are marked in term of time with a reasonable consumption of hardware resources. Smart grids are used as a real case study to apply the proposed parallel verification methodology. As perspectives, the proposed contribution can be applied to other formal tools that support the reconfigurable hierarchical design and a comparison between related performances can be discussed.

REFERENCES

Alur R. and Yannakakis M., 2001. *Model checking of hierarchical state machines. ACM Transactions on Programming Languages and Systems (TOPLAS)*, 23, no. 3, 273–303.

Barnat J.; Brim L.; Ceska M.; and Rockai P., 2010. *Divine: Parallel distributed model checker.* In *2010 Ninth International Workshop on Parallel and Distributed Methods in Verification, and Second International Workshop on High Performance Computational Systems Biology*. IEEE, 4–7.

Bengtsson J. and Yi W., 2003. *Timed automata: Semantics, algorithms and tools.* In *Advanced Course on Petri Nets*. Springer, 87–124.

Boucheneb H.; Gardey G.; and Roux O.H., 2009. *TCTL model checking of time Petri nets. Journal of Logic and Computation*, 19, no. 6, 1509–1540.

Bouyer P. et al., 2008. *Model checking timed automata.*

Jiang Z.; Khalgui M.; Mosbahi O.; and Jaouadi A., 2014. *A novel hierarchical multi-agent architecture for automatic restoration of smart grids. International Journal of Control and Automation*, 7, no. 1, 153–170.

La Torre S.; Napoli M.; Parente M.; and Parlato G., 2008. *Verification of scope-dependent hierarchical state machines. Information and Computation*, 206, no. 9-10, 1161–1177.

Lin H. and Li P., 2014. *Parallel hierarchical reachability analysis for analog verification.* In *Proceedings of the 51st Annual Design Automation Conference*. ACM, 1–6.

Ouaknine J. and Worrell J., 2008. *Some recent results in metric temporal logic.* In *International Conference on Formal Modeling and Analysis of Timed Systems*. Springer, 1–13.

Roufaida B.; Laid K.; Mohamed K.; and Zhiwu L., 2019. *Reconfigurable Hierarchical Timed Automata: Modeling and Stochastic Verification.* In *2019 IEEE International Conference on Systems, Man, and Cybernetics*. IEEE.

Sklyarov V. and Skliarova I., 2008. *Design and implementation of parallel hierarchical finite state machines.* In *2008 Second International Conference on Communications and Electronics*. IEEE, 33–38.

SIMULATION USING FUZZY ALGORITHMS

ON AGGREGATION OF GENERALIZED DISTANCES AND AN APPLICATION TO ANALYSIS OF ALGORITHMS

José Guerrero
Universitat de les Illes Balears
Cra. Valldemossa km. 7.5
07122-Palma de Mallorca, Spain
E-mail: jose.guerrero@uib.es

Juan-José Miñana
Universitat de les Illes Balears,
Cra. Valldemossa km. 7.5
07122-Palma de Mallorca, Spain
E-mail: jj.minana@uib.es

Oscar Valero
Universitat de les Illes Balears,
Cra. Valldemossa km. 7.5
07122-Palma de Mallorca, Spain
E-mail: o.valero@uib.es

KEYWORDS
partial quasi-metric, complexity analysis of algorithms, denotational semantics, aggregation.

ABSTRACT

Generalized dissimilarities have been used successfully as efficient tools in modeling in Computer Science. In fact, partial metrics and quasi-metrics have shown to be useful in the study of recursive denotational specifications for programming languages and to discuss the complexity analysis of algorithms, respectively. The aforesaid models were possible thanks to obtaining quantitative fixed point methods for such dissimilarities. However, the aforesaid quantitative fixed point techniques are developed separately without any relationship between them. In this short note we want to explicitly state that both techniques could be unified via aggregation in order to develop a framework which remains valid at the same time for modeling in denotational semantics and in complexity analysis of algorithms.

INTRODUCTION

In applied sciences, the problem of merging a collection of data (inputs) into a single one datum (output), which contains information on each of the inputs, plays a central role. Typical fields in which this kind of problem arises in a natural way are robotics, decision making, image processing, medical diagnosis, machine learning, pattern recognition, econometrics or business management. Most of the time, in the indicated fields, the information is coded as numerical data and, therefore, such data must be fused in order to obtain a unique numerical value that helps us make a working decision. Many methods to merge these numerical inputs are based on the so-called aggregation functions. A good monograph for a recent account of such class of functions are (Beliakov et al. 2016). Sometimes the nature of the problem imposes that the aggregation method provides an output preserving the fundamental properties of the inputs. This is the case when a collection of dissimilarities are merged in order to obtain a new one which represents any type of global dissimilarity and whose dissimilarity values allow us to make a decision. The utility of dissimilarities in modeling problems in applied sciences has motivated the study of those functions that allow to merge a collection of generalized dissimilarities into a new one. Thus, a characterization of those functions that merge a collection of metrics into a new one was given in (Borsík and Doboš 1981). A general solution to the problem of merging S-metrics and pseudometrics was provided in (Pradera et al. 2000; Pradera et al. 2002; Pradera and Trillas 2002). Several linkages between the aggregation operators theory and metric aggregation functions have been given in (Casasnovas and Rosselló 2005) and (Yager 2010). In (Mayor and Valero 2010), the original work of Borsík and Doboš was extended to the framework of quasi-metrics. The aggregation problem for partial metrics and, in addition, relationships between this problem and the (quasi-)metric one were explored in (Massanet and Valero 2012). Recently, a refinement of the original Borsík and Doboš characterization has been yielded in (Mayor and Valero 2019).

Inspired, on the one hand, by the interest that causes the dissimilarities aggregation problem and, on the other hand, by the utility of generalized dissimilarities in applied sciences, we will focus our attention on the problem of developing a framework which remains valid for modeling in denotational semantics and in complexity analysis of algorithms at the same time via the aggregation of a quasi-metric and a partial metric.

THE NEED FOR AGGREGATION IN ANALYSIS OF ALGORITHMS

As exposed before, partial metrics have been applied successfully to denotational semantics and quasi-metrics have been used in asymptotic complexity analysis of algorithms.

In denotational semantics one of the targets is to verify the correctness of recursive algorithms through mathematical models. In many programming languages, recursive algorithms can be constructed by means of procedures in such a way that the meaning of such a procedure is expressed in terms of its own meaning. A typical, and illustrative, example is the procedure which computes the factorial function. The aforesaid procedure when computes the factorial of a positive integer number uses the following recursive denotational specification:

$$fact(n) = \begin{cases} 1 & \text{if } n = 1 \\ nfact(n-1) & \text{if } n > 1 \end{cases} . \qquad (1)$$

When discussing whether a recursive denotational specification of a procedure is meaningful, fixed point mathematical techniques are used in such a way that the meaning of such recursive denotational specification is obtained as the fixed point of a non-recursive mapping associated to the denotational specification. In the particular case of the factorial function the aforesaid non-recursive mapping ϕ_{fact} will be given as follows:

$$\phi_{fact}(f)(n) = \begin{cases} 1 & \text{if } n = 1 \\ nf(n-1) & \text{if } n > 1 \text{ and } n-1 \in \text{dom} f \end{cases}, \quad (2)$$

where ϕ_{fact} is acting over the set of partial functions. Of course, the entire factorial function is given by the unique fixed point of the non-recursive mapping ϕ_{fact}. For a detailed treatment of the set of partial functions and its applications to denotational semantics we refer the reader to (Davey and Priestley 1990).

In order to prove the existence of a fixed point of the non-recursive mapping ϕ_{fact} (given by (2)), Matthews introduced in (Matthews 1994) the so-called Baire partial metric space. Such a space consists of the pair (Σ_∞, p_B), where Σ_∞ is the set of finite and infinite sequences over a non-empty alphabet Σ and the partial metric p_B is given by $p_B(v,w) = 2^{-l(v,w)}$ for all $x, y \in \Sigma_\infty$ with $l(v,w)$ denoting the the longest common prefix of the words v and w when it exists and $l(v,w) = 0$ otherwise. Of course the convention that $2^{-\infty} = 0$ is adopted. Notice that every partial function f can be identified with a word $w^f \in \mathbb{N}_\infty$ such that $w^f = w_1^f w_2^f \dots w_k^f$ with $\text{dom} f = \{1, \dots, k\}$ and $w_i^f = f(i)$ for all $i \in \text{dom} f$. Then it can be verified that ϕ_{fact} is a contraction from $(\mathbb{N}_\infty, p_B)$ into itself with $\frac{1}{2}$ as contractive constant.

Usually the running time of computing of the recursive algorithm that performs the computation of the meaning of a recursive denotational specification is analyzed in conjunction with the correctness of such a recursive denotational specification. In this direction, it was introduced the so-called complexity space which consists of the pair $(\mathcal{C}, q_{\mathcal{C}})$, where

$$\mathcal{C} = \{f : \mathbb{N} \longrightarrow (0, \infty] : \sum_{n=1}^{\infty} 2^{-n} \frac{1}{f(n)} < \infty\}$$

and $q_{\mathcal{C}}$ is the quasi-metric on $\mathcal{C}$ defined by

$$q_{\mathcal{C}}(f,g) = \sum_{n=1}^{\infty} 2^{-n} \max\left(\frac{1}{g(n)} - \frac{1}{f(n)}, 0\right).$$

Clearly the convention that $\frac{1}{\infty} = 0$ is adopted (see (Schellekens 1995)).

The running time of computing of an algorithm can be associated to a function belonging to $\mathcal{C}$. In addition, the numerical value $q_{\mathcal{C}}(f,g)$ (the complexity distance from f to g) can be interpreted as the relative progress made in lowering the complexity by replacing any program P with complexity function f by any program Q with complexity function g. Moreover, the condition $q_{\mathcal{C}}(f,g) = 0$ can be understood as f is "at least as efficient" as g on all inputs. Notice that $q_{\mathcal{C}}(f,g) = 0$ implies that $f(n) \leq g(n)$ for all $n \in \mathbb{N}$. The last fact is crucial when the asymptotic upper bound of the complexity of an algorithm wants to be specified. In fact, $q_{\mathcal{C}}(f,g) = 0$ implies that $f \in \mathcal{O}(g)$, where $f \in \mathcal{O}(g)$ means that there exist $n_0 \in \mathbb{N}$ and $c \in \mathbb{R}_+$ such that $f(n) \leq cg(n)$ for all $n \in \mathbb{N}$ with $n \geq n_0$. Furthermore, it must be pointed out that the asymmetry of dissimilarity $q_{\mathcal{C}}$ is key when one wants to give information about the growth in complexity when a program is replaced by another one. A metric would be able to be used in order to provide information on the growth but, nevertheless, it could not yield which program is more efficient.

The running time of computing of an algorithm that computes the factorial of a nonnegative integer number, through the recursive denotational specification (1), is the solution to the following recurrence equation

$$T_{fact}(n) = \begin{cases} c & \text{if } n = 1 \\ T_{fact}(n-1) + c & \text{if } n > 1 \end{cases}, \quad (3)$$

where $c \in \mathbb{R}_+$ ($c > 0$) is the time taken by the algorithm to obtain the solution to the problem on the base case. Schellekens' approach provides an upper bound of the solution to the recurrence equation (3) making the following statement. Consider the subset $\mathcal{C}_c$ of $\mathcal{C}$ given by $\mathcal{C}_c = \{f \in \mathcal{C} : f(1) = c\}$. Define the mapping $G_{fact} : \mathcal{C}_c \to \mathcal{C}_c$ by

$$G_{fact}(f)(n) = \begin{cases} c & \text{if } n = 1 \\ f(n-1) + c & \text{if } n \geq 2 \end{cases} \quad (4)$$

for all $f \in \mathcal{C}_c$. Clearly $f \in \mathcal{C}_c$ is the unique solution to the recurrence equation (3) if and only if f is the unique fixed point of the mapping G_{fact}. The uniqueness of fixed point warranties that there exists only one function belonging to $\mathcal{C}_c$ that can be identified with the running time of computing of the algorithm under consideration. Moreover, if $f_{G_{fact}}$ denotes the fixed point of G_{fact}, then $f_{G_{fact}} \in \mathcal{O}(g)$ provided that there exists $g \in \mathcal{C}_c$ with $q_{\mathcal{C}}|_{\mathcal{C}_c}(G_{fact}(g), g) = 0$ and, this way, allows us to yield an asymptotic upper bound of the running time of computing of the algorithm computing the factorial.

Observe that in order to analyze, on the one hand, the meaning of a recursive denotational specification and, on the other hand, the running time of computing of the algorithm which performs the computation of such a meaning, the exposed quantitative techniques are independent used without any relationship stablished between both. Note that one of them is based on the use of partial metrics and the other one is based on the use of quasi-metrics.

At first glance it seems difficult to combine two approaches so that we can build a unique framework which allows us to carry out formally, via fixed point methods, the two tasks at the same time. However, it seems reasonable that we consider a new structure which arise merging the both original ones in such a way that the main properties coming from both different nature inputs are kept and the original dissimilarities (a partial metric and a quasi-metric) are fused in a global

dissimilarity. Taking this into account and the fact that aggregation functions become of great importance because they allow to develop efficient numerical fusion methods, it would be suitable to involve such functions to define a new type of generalized dissimilarity $\Phi(p_B(v,w), q_{\mathcal{C}}|_{\mathcal{C}_c}(f,g))$ for all $v, w \in \mathbb{N}_\infty$ and $f, g \in \mathcal{C}_c$, which might not be either a partial metric or a quasi-metric, that could help us to obtain a general fixed point result for contractions obtained through aggregation and, thus, to develop quantitative fixed point methods which remains valid to discuss simultaneously the complexity analysis of algorithms and the meaning of recursive denotational specifications for programming languages. Therefore, it would be very positive that multidisciplinary research groups (formed by Mathematicians and Engineers) focus their efforts on combining generalized metric structures and information fusion to show the way in which the aforementioned target can be achieved.

ACKNOWLEDGEMENTS

The authors acknowledge the financial support from FEDER/Ministerio de Ciencia, Innovación y Universidades-Agencia Estatal de Investigación/_Proyecto PGC2018-095709-B-C21. This work is also partially supported by Programa Operatiu FEDER 2014-2020 de les Illes Balears, by project PROCOE/4/2017 (Direcció General d'Innovació i Recerca, Govern de les Illes Balears) and by project ROBINS. The latter has received research funding from the EU H2020 framework under GA 779776. This publication reflects only the authors views and the European Union is not liable for any use that may be made of the information contained therein.

REFERENCES

G. Beliakov, H. Bustince and T. Calvo-Sánchez, A Practical Guide to Averaging Functions, Springer, 2016.

J. Borsík, J. Doboš, On a product of metric spaces, Mathematica Slovaka 31 (1981), 193-205.

J. Casasnovas, F. Rosselló, Averaging fuzzy biopolymers, Fuzzy Sets and Systems 152 (2005), 139-158.

B.A. Davey, H.A. Priestley, Introduction to Lattices and Order, Cambridge University Press, 1990.

M.M. Deza, E. Deza, Encyclopedia of Distances, Springer-Verlag, 2009.

S. Massanet, O. Valero, New results on metrics aggregation, In: Proc. 16th Spanish Conference on Fuzzy Technology and Fuzzy Logic, European Society for Fuzzy Logic and Technology (Edited by G.I. Sainz-Palmero et al.), pp. 558-563. European Society for Fuzzy Logic and Technology. Valladolid (2012).

G. Mayor, O. Valero, Metric aggregation functions revisited, European Journal of Combinatorics 80 (2019), 390-400.

G. Mayor, O. Valero, Aggregation of asymmetric distances in Computer Science, Information Sciences 180 (2010), 803-812

S.G. Matthews, Partial Metric Topology, Annals of the New York Academy of Sciences 728 (1994), 183-197.

A. Pradera, E. Trillas, E. Castiñeira, On distances aggregation, In: Proc. Information Processing and Management of Uncertainty in Knowledge-Based Systems International Conference IPMU 2000, (Edited by B. Bouchon-Meunier et al.), Vol. II, pp. 693-700. Universidad Politécnica de Madrid Press, Madrid (2000).

A. Pradera, E. Trillas, E. Castiñeira, On the aggregation of some classes of fuzzy relations, In: Technologies for Constructing Intelligent Systems (Edited by B. Bouchon-Meunier et al.), pp. 126-136. Springer-Verlag, (2002).

A. Pradera, E. Trillas, A note on pseudometrics aggregation, Intertanional Journal of General Systems 31 (2002), 41-51.

M.P. Schellekens, The Smyth completion: a common foundation for denotational semantics and complexity analysis, Electronic Notes in Theoretical Computer Science 1 (1995), 211-232.

R.R. Yager, Norms induced from OWA operators, IEEE Transactions on Fuzzy Systems 18 (2010), 57-65.

BIOGRAPHY

JOSÉ GUERRERO received his degree in Computer Science from the University of the Balearic Islands (UIB) where he is currently a lecturer and postdoctoral researcher. His research interest includes Multi-robot Task Allocation Mechanisms with Auction and Swarm Coordination Mechanisms and Mathematical Models applied to Computer Science and Artificial Intelligence.

JUAN JOSÉ MIÑANA is Associate Professor (tenure track) at Universitat de les Illes Balears (Departament de Ciències Matemàtiques i Informàtica). He received his M.Sc. in Mathematics from Universitat de València in 2010 and, in 2015, he obtained a Ph.D. in Mathematics from Universitat Politècnica de València. His research activity focus on Topology and Applied Mathematics. Concretely, on the study of generalized metrics and its applications.

OSCAR VALERO is Associate Professor at Universitat de les Illes Balears (Departament de Ciències Matemàtiques i Informàtica). He received his M.Sc. in Mathematics from Universitat de València in 2000. Later on, in 2003, he obtained a Ph.D. in Mathematics from Universitat Politècnica de València in the Doctoral Program of Multidisciplinary Mathematics. His research activity focus on Mathematical Models applied to Computer Science, Artificial Intelligence, Engineering, Economics and Medicine.

A MODIFICATION OF OWD AGGREGATION OPERATOR AND ITS APPLICATION TO K-MEANS ALGORITHM

Juan-José Miñana
Universitat de les Illes Balears,
Ctra. Valldemossa km. 7.5
07122-Palma de Mallorca, Spain
E-mail: jj.minana@uib.es

Mateu Morro
Hotelbeds Group
Cami de Son Fangos,100 Block Edif. Mirall, Torre A.
07007-Palma de Mallorca, Spain
E-mail: mmorro@hotelbeds.com

Oscar Valero
Universitat de les Illes Balears,
Cra. Valldemossa km. 7.5
07122-Palma de Mallorca, Spain
E-mail: o.valero@uib.es

KEYWORDS
aggregation, OWA operator, pattern recognition, clustering, K-Means

ABSTRACT

Aggregation functions have been deeply studied in the literature and many applications to real life problems have been obtained. A specific aggregation function, known as OWA operator, has shown to be specially useful in decision making problems. In this paper we show that OWAs operators acting on a collection of distances can be considered as an appropriate dissimilarity measure in cluster analysis. Concretely we introduce a new dissimilarity measure based on OWAs and we show that it is able to overcome a drawback that presents the classic K-Means algorithm when the Euclidean distance is used as a dissimilarity between objects. A collection of experiments are presented showing that our new approach outperforms the classic one.

INTRODUCTION

The fusion of information appears, in a natural way, in many fields as mathematics, computer science, engineering, economics, finance, among others. It consists in merging a set of data, given as inputs, to obtain a single one datum as an output, in such a way that the output represents or contains most faithfully the information of all inputs.
The study of how to combine the information provided by the inputs is essential in order to get the adequate output. The theory of aggregation functions tackles such an study when all considered inputs are real values.
Recall that, given $a, b \in [-\infty, \infty]$ with $a < b$, an *aggregation function* is a function $A : [a, b]^n \to [a, b]$ that is non-decreasing (i.e. $A(x_1, \ldots, x_n) \leq A(y_1, \ldots, y_n)$ whenever $x_i \leq y_i$ for each $i \in \{1, \ldots, n\}$) and satisfying the boundary conditions (i.e $A(a, \ldots, a) = a$ and $A(b, \ldots, b) = b$). A well-known example of aggregation function is the so-called arithmetic mean.

Aggregation functions have been deeply studied in the literature (see (Grabisch et al. 2009) and references therein), and their range of applicability is so wide in the aforementioned fields. Concretely, aggregation functions provide a useful tool in decision making in such a way that one has an overall of the decision (the output) under some given characteristics (the inputs). It can be described as follows: let $A : [a, b]^n \to [a, b]$ be an aggregation function and consider, for each $i \in \{1, \ldots, n\}$ the coordinate function A_i. Such coordinate functions are interpreted as the criteria taken into account in order to make a decision. Let $X \subseteq [a, b]^n$ be a non-empty set of alternatives. For each $x \in X$, the value $A_i(x) \in [a, b]$ can be interpreted as the degree to which x satisfies that criterion. Then the aggregation function A can be understood as a tool to yield a global scoring which quantifies the degree to which an alternative x satisfies at the same time all criteria under consideration.
Based on such an application of aggregation functions, in (Yager 1988) the notion of Ordered Weighted Averaging (briefly OWA) aggregation operator was introduced. Let us recall that, given a (weighting) vector $W \in [0, 1]^n$ such that $\sum_{i=1}^n w_i = 1$, an *OWA aggregation operator* of dimension n is an aggregation function $OWA : [a, b]^n \to [a, b]$ such that

$$\text{OWA}(x_1, \ldots, x_n) = \sum_{i=1}^{n} w_i x_{(i)},$$

where $x_{(i)}$ denotes the ith largest element in the collection $\{x_1, \ldots, x_n\}$.
Based on the notion of OWA aggregation operator, in (Xu and Chen 2008) introduced the concept of Ordered Weighted Distance (briefly OWD) measure. Attending to (Xu and Chen 2008), given a (weighting) vector $W \in [0, 1]^n$ such that $\sum_{i=1}^n w_i = 1$, an *OWD measure* of dimension n is a function $\text{OWD} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ satisfying the following condition:

$$\text{OWD}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} w_i |x_{(i)} - y_{(i)}|,$$

where $|x_{(i)} - y_{(i)}|$ denotes the ith largest element in the collection $\{|x_1 - y_1|, \ldots, |x_n - y_n|\}$. Notice that an OWD measure is just an OWA operator, induced by a weighting vector W, applied to the collection of distances $\{|x_1 - y_1|, \ldots, |x_n - y_n|\}$. Since the notion of OWD measure was introduced many applications to decision making in insurance management, investment management or human resources management have been obtained in (Merigó and Casanovas 2008), (Merigó and Gil-Lafuente 2008), (Merigó and Casanovas 2010), (Merigó and Casanovas 2010b), (Merigó and Casanovas 2011), (Merigó and Casanovas 2011b) and (Casanovas et al. 2016).

Related to the study of OWAs operators and OWD measures, in (Mayor and Valero 2010) a depth study about how a collection of distances can be merge by means of a function was made. Concretely, it was shown the utility of OWAs operators to this target and, in addition, that the celebrated Euclidean distance can be understood as the aggregation of a collection of distances.

In many clustering analysis algorithms, the dissimilarity between objects to be sorted out is measured by means of the Euclidean distance (Theodoridis and Koutroumbas 2006). A typical example of this kind of algorithms is the so-called k-means algorithm. However, we have detected a drawback when the Euclidean distance is used in the classic K-Means algorithm. Inspired, on the one hand, by the applicability of the OWD measures and by the fact that the Euclidean distance can be understood as the result of an aggregation operator acting on a collection of distances and, on the other hand, by the inconvenience found, in this paper we propose a modification of the notion of OWD measure, that we call OWDr measure, to overcome the aforesaid drawback. Moreover, we present a modified k-means algorithm which is based on the use of an OWDr measure instead of the Euclidean distance. The new algorithm is implemented and a collection of experiment, all carried out by means of RSTUDIO, are presented when the world-renowned Iris dataset is taken under consideration. The obtained results show that our new approach outperforms the classic one.

CLUSTER ANALYSIS AND K-MEANS ALGORITHM

Cluster analysis consists in organizing a collection of objects or patterns, usually vectors of n components, into different groups, known as clusters, based on similarity. Such an organization should fulfill that patterns in the same cluster are more similar to each other than those contained in different clusters.

Clustering plays a central role in data analysis used in different fields as pattern recognition, image segmentation, machine learning, bioinformatics, among others.

Due its wide range of applicability, the problem of clustering has been extensively tackled in the literature (see (Theodoridis and Koutroumbas 2006; Rajamaran et al. 2012) and references therein) and several clustering algorithms have been proposed.

Following (Jain 2010) such algorithms can be divided in two families: hierarchical and partitional. On the one hand, hierarchical algorithms find recursively nested clusters following two different methods. In one of them, the process is started with each data point in its own cluster and then, the most similar pair of clusters are merged to form a cluster hierarchy. In the other one, the process is started with all the data points in the same cluster and then, each cluster is recursively divided into smaller clusters. On the other hand, partitional clustering algorithms find all the clusters simultaneously as a partition of the data and do not impose a hierarchical structure (see (Theodoridis and Koutroumbas 2006) for a deeper treatment). The most popular partitional algorithm of clustering is the so-called k-means, and although it was proposed over 50 years ago, nowadays it is one of the most widely used algorithms for clustering due to its easy implementation, simplicity, efficiency, and empirical success. Below, we detail the essence of such an algorithm.

Consider $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ a set of n d-dimensional points to be separated in k different clusters $\{C_1, \ldots, C_k\}$. The aim of k-means consists in obtaining a partition such that the minimum square error between the empirical mean (cluster center) of a cluster and the points in the cluster is minimized. The procedure runs as follows.

1. Select an initial partition with k clusters and fix k initial centroids (cluster centers) $\mathbf{c}_1^0, \ldots, \mathbf{c}_k^0$.
2. For each step $t \geq 0$, every $\mathbf{x}_i \in X$ is located in the cluster C_l^t such that the distance with the centroid $\mathbf{c}_l^t$ is minimum, for all $l \in \{1, \ldots, k\}$.
3. The centroid of each cluster is recalculated for the next step $t + 1$ by calculating the arithmetic mean of each cluster.
4. If a convergence criterion is not met (cluster membership stabilizes), go to step 2.

On account of above procedure, one can observe that K-means algorithm depends on the number of clusters k (initial selected centroids) and the distance used to measure the difference between the centroid of a cluster and each point of X to be assigned into a cluster. Different works have contributed to the study of how the centroids can be chosen and what is the appropriate distance to measure the similarities (see (Jain 2010; Jain el al. 1999) and references therein). In this paper we approach the last one, i.e., how to measure the difference between a centroid and a point $\mathbf{x}_i \in X$. Usually, such a measurement is made by means of the Euclidean distance. However, we have detected a handicap when the aforementioned distance is under consideration and the K-means algorithm is applied to the Iris dataset. So we propose a modification of K-means in such a way that dissimilarities are quantified through a new OWD measure instead of the Euclidean distance. In Section "New OWD Measure" we motivate our new approach and propose the new measure to overcome the aforesaid inconvenience. Later on, in Section "Experimental Results" we illustrate its usefulness in cluster

analysis when it is applied to Iris dataset and, in addition, we show that our new approach outperforms the classic one.

NEW OWD MEASURE

In this section we provide a modification of the notion of OWD measure to improve the K-means algorithm when a concrete type of dataset is analyzed. We are focused on those datasets whose elements are described by a vector $x \in \mathbb{R}_+^n$ such that each x_i is expressed at different scales, where $\mathbb{R}_+ = [0, \infty[$. As exposed before, the Euclidean distance is used usually to compare the centroids and the elements to be clustered on the K-means algorithm. Nevertheless, this way of proceeding does not provide a measurement that takes into account the different scales of the components and this could produce an inconvenience in the assignment of the different elements (patterns) to the closest cluster center. Indeed, below, we illustrate, how the Euclidean distance could provide an erroneous assignment. To this end, consider a dataset X of 2-dimensional points. Assume, without loss of generality, that in any step of our clustering precess we have to assign a point to its cluster center when we have the followoing two centroids: $\mathbf{c}_1 = (100, 0.001)$ and $\mathbf{c}_2 = (90, 2)$. If we try to measure, by means of the Euclidean distance, the difference from the point $\mathbf{x} = (97, 1.999)$ to the centroids c_1 and c_2 of the clusters C_1 and C_2, respectively, then we obtain that the closest cluster center is c_1 and, thus, $\mathbf{x}$ will be included in C_1, since

$$d_e(\mathbf{c}_1, \mathbf{x}) = \sqrt{(100-97)^2 + (0.001-1.999)^2} \approx 3.6 <$$

$$< 7 \approx \sqrt{(90-97)^2 + (2-1.999)^2} = d_e(\mathbf{c}_2, \mathbf{x}).$$

Nevertheless, sometimes it could be more appropriate to assign $\mathbf{x}$ to the cluster C_2. An example of this situation could occur in those cases in which the second component of the elements of X takes values just in $[0, 2]$ and, at the same time, the first component takes values in the whole $[0, 100]$. In such a case, the difference between the second components of $\mathbf{x}$ and $\mathbf{c}_1$ is almost the maximum that can be gotten in the range in which they take their values. So, the second component of this two elements differs almost the maximum and, thus, it seems natural to assign the pattern $\mathbf{x}$ to cluster C_2. However, the Euclidean distance classifies them at the same cluster. Of course, this fact is a disadvantage that could have great consequence in the final clustering. Observe that the l_1 distance, another important class of distance used in clustering (see (Theodoridis and Koutroumbas 2006)), will also assign $\mathbf{x}$ to cluster C_1 and so l_1 does not overcome the aforesaid inconvenient.

Inspired by the preceding drawback, we propose a new dissimilarity measure which computes a relative difference in each component separately between the pattern and the centroid and then, such a collection of differences is aggregated by means of an OWA operator. So, we propose a measure which is able to provide a global difference between centroid and pattern, in the spirit of the Euclidean distance, but comparing coordinate-to-coordinate and, thus, being sensitive to the difference in scales of all components involved in the measurement.

Based on the exposed above, we propose the following modification of the concept of OWD measure which will be crucial in our subsequent discussion.

Definition 1. An Ordered Weight Distance relative (briefly OWDr) measure of dimension n is a function OWDr : $\mathbb{R}_+^n \times \mathbb{R}_+^n \to \mathbb{R}_+^n$ that has associated a weighting vector $W \in [0,1]^n$ with $\sum_{i=1}^n w_i = 1$, satisfying the following condition:

$$\text{OWDr}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} w_i d_r(x_{(i)}, y_{(i)}),$$

where $d_r(x_{(i)}, y_{(i)})$ denotes the ith largest element in the collection of distances $\{d_r(x_1, y_1), \ldots, d_r(x_n, y_n)\}$, and $d_r : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ is the distance given by

$$d_r(x, y) = \begin{cases} 0 & \text{x=y=0} \\ \frac{|x-y|}{\max\{x,y\}} & otherwise \end{cases}.$$

Next we show that an OWDr measure is able to overcome the inconvenient above exposed.

Consider the weighting vector $W = (2/3, 1/3)$, the pattern $\mathbf{x} = (97, 1.999)$ and the centroids $\mathbf{c}_1 = (100, 0.001)$, $\mathbf{c}_2 = (90, 2)$. On the one hand, $\frac{|x_1 - c_{1,1}|}{\max\{x_1, c_{1,1}\}} = \frac{3}{100} = 0.03$ and $\frac{|x_2 - c_{1,2}|}{\max\{x_2, c_{1,2}\}} = \frac{1.998}{2} = 0.999$. Then,

$$\text{OWDr}(\mathbf{x}, \mathbf{c}_1) = 2/3 \cdot 0.999 + 1/3 \cdot 0.03 = 0.676.$$

On the other hand, $\frac{|x_1 - c_{2,1}|}{\max\{x_1, c_{2,1}\}} = \frac{7}{97} = 0.072165$ and $\frac{|x_2 - c_{2,2}|}{\max\{x_2, c_{2,2}\}} = \frac{0.001}{2} = 0.0005$. Then,

$$\text{OWDr}(\mathbf{x}, \mathbf{c}_2) = 2/3 \cdot 0.072165 + 1/3 \cdot 0.0005 = 0.048277.$$

Therefore, $\mathbf{x}$ will be assigned to the cluster C_2, contrary to the Euclidean distance case, as one would expect.

EXPERIMENTAL RESULTS

After introducing the notion of OWDr measure, in this section we show its usefulness in cluster analysis. To this end, we check this new measurement for k-means algorithm in the well-known Iris dataset with 150 flower images corresponding to Versicolor, Virginica and Setosa species.

All the experiments have been carried out using RSTUDIO. The functions *kmeans()* of the library *ppclust*, *factoextra* and *cluster* have been used to obtain the results of the classic K-Means. The entries of the functions have been the Iris dataset (removing the variable that classifies the plants) and the number of clusters, which has been fixed as $k = 3$.

The proposed modification of K-Means has been implemented following the classic procedure explained in Section "Cluster Analysis and K-Means Algorithm" modifying the step 2, where an OWDr measure has been used to calculate

the dissimilarities instead of the Euclidean distance. From now one, we will refer to the new algorithm as OWDr K-Means. The experiments have converged using a number of iterations lower than the maximum one fixed as additional convergence criterion.

Figure 3 shows the results obtained in the experiments where the centroids of the first clusters fixed in step 1 of the algorithm have been randomly assigned in all cases. Moreover, the results have been obtained using the same initial centroids when running the classic K-Means and OWDr K-Means. It must be stressed, that when running OWDr K-Means many weighting vectors have been tested. However we show a few one, since the obtained results mainly match up with one of those shown in Figure 3. Moreover, in all experiments the accuracy rate for the classic algorithm does not exceed 86% while that the accuracy rate for the OWDr K-Means is between 94% and 96%. Clearly, Figure 3 shows that in the best case the classic K-Means fails in 16 flowers. Nevertheless, the OWDr K-Means fails only in 9 flowers in the worst case. The best result is obtained when the weighting vector is fixed as either $W = [0.4, 0.3, 0.2, 0.1]$ or $W = [0.7, 0.2, 0.1, 0]$. When the weighting W is considered as $W = [0.25, 0.25, 0.25, 0.25]$, the accuracy decreases, although it still improves the classic K-means.

Figures 1 and 2 show visually the obtained results in the best cases for classic K-Means and for OWDr K-Means. We have used the library *tidyverse* to generate such figures. Notice that OWDr K-Means, comparing coordinate-to-coordinate, overcomes successfully, contrary to the classic algorithm, the difficulty of differentiating the Versicolor and Virginica species. Moreover, both algorithms classify Setosa species perfectly.

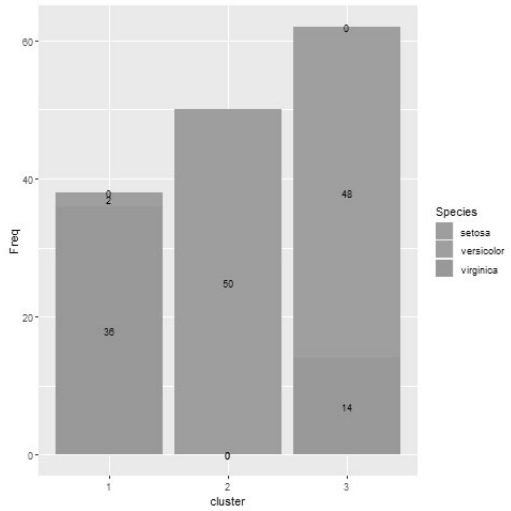


Figure 1: Iris classical K-Means: results for test 1

CONCLUSIONS AND FUTURE WORK

In this paper we have shown that OWAs operators acting on a collection of distances can be considered as an appropriate

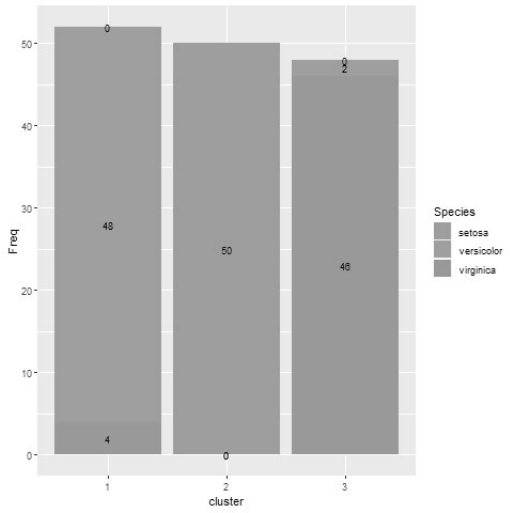


Figure 2: Iris OWDr K-Means: results for the OWDr measure induced by the weighting vector $(0.4, 0.3, 0.2, 0.1)$.

	Correct Clustering	Incorrect Clustering	Accuracy Rate
Classic K-Means (Test 1)	134	16	89.33%
Classic K-Means (Test 2)	129	21	86%
Classic K-Means (Test 3)	134	16	89.33%
OWDr K-Means W=(0.25, 0.25, 0.25, 0.25)	141	9	94%
OWDr K-Means W=(0.3, 0.3, 0.2, 0.2)	142	8	94.67%
OWDr K-Means W=(0.4, 0.3, 0.2, 0.1)	144	6	96%
OWDr K-Means W=(0.7, 0.2, 0.1, 0)	144	6	96%

Figure 3: Results obtained for classic K-Means and OWDr K-Means with different weighting vectors

dissimilarity measure in cluster analysis. Concretely we have introduced a new dissimilarity measure based on OWAs and we have shown that it is able to overcome a drawback that presents the classic K-Means algorithm when the Euclidean distance is used as a dissimilarity between objects. A collection of experiments have been presented showing that our new approach outperforms the classic one. In all aforesaid experiments a weighting vector has been involve. However, the weighting vector selection has been made in a heuristic way. As a future way we will investigate the issue of generating by means of a procedure the weighting vector inducing the OWA operator that best fits to the dataset under consideration. A large number of experiments in order to compare the new possible results with those provided by the classic K-Means algorithm will be performed.

ACKNOWLEDGEMENTS

J.J. Miñana and O. Valero acknowledge financial support from FEDER/Ministerio de Ciencia, Innovación y Universidades-Agencia Estatal de Investigación/_Proyecto PGC2018-095709-B-C21. This work is also partially supported by Programa Operatiu FEDER 2014-2020 de les

Illes Balears, by project PROCOE/4/2017 (Direcció General d'Innovació i Recerca, Govern de les Illes Balears) and by project ROBINS. The latter has received research funding from the EU H2020 framework under GA 779776. This publication reflects only the authors views and the European Union is not liable for any use that may be made of the information contained therein.

REFERENCES

M. Casanovas, A. Torres-Martínez, J.M. Merigó, Decision making in reinsurance with induced OWA operators and Minkowski distances Cybernetics and Systems 47:6 (2016), 460-477.

M. Grabisch, J.L. Marichal, R. Mesiar, E. Pap, Aggregation Functions, Cambridge University Press, 2009.

A.K. Jain, Data clustering: 50 years beyond K-means, Pattern Recognition Letters 31 (2010), 651-666.

A.K. Jain, M.N. Murty, P.J. Flynn, Data clustering: a review, ACM Computing Surveys (CSUR) 31:3 (1999), 264-323.

G. Mayor, O. Valero, Aggregation of asymmetric distances in Computer Science, Information Sciences 180 (2010), 803-812.

J.M. Merigó, A.M. Gil-Lafuente, Using the OWA operator in the Minkowski distance, International Journal of Computer Science 3:3 (2008), 149-157.

J.M. Merigó, M. Casanovas, The induced Minkowski ordered weighted averaging distance operator, in: Proceedings of XIV Congreso Español sobre Tecnologías y Lógica fuzzy ESTYLF 2008, pp. 35-641, European Centre for Soft Computing, (2008).

J.M. Merigó, M. Casanovas, Decision making with distance measures and linguistic aggregation operators, International Journal of Fuzzy Systems 12:3 (2010), 190-198.

J.M. Merigó, M. Casanovas, Induced and heavy aggregation operators with distance measures, Journal of Systems Engineering and Electronics 21:3 (2010), 431-439.

J.M. Merigó, M. Casanovas, Decision-making with distance measures and induced aggregation operators, Computers and Industrial Engineering 60 (2011), 66-76.

J.M. Merigó, M. Casanovas, A new Minkowski distance based on induced aggregation operators, International Journal of Computational Intelligence Systems 4 (2011), 123-33.

A. Rajaraman, J. Leskovec, J. David Ullman, Mining of Massive Datasets, Cambridge University Press, 2012.

S. Theodoridis, K. Koutroumbas, Pattern Recognition, Elsevier Academic press, USA, 2006.

R.R. Yager, On ordered weighted averaging aggregation operators in multicriteria decision making, IEEE Transactions on Systems, Man, and Cybernetics 18:1 (1988), 183-190.

Z. Xu, J. Chen, Ordered weighted distance measure, Journal of Systems Science and Systems Engineering 17:4 (2008), 432-445.

JUAN JOSÉ MIÑANA is Associate Professor (tenure track) at Universitat de les Illes Balears (Departament de Ciències Matemàtiques i Informàtica). He received his M.Sc. in Mathematics from Universitat de València in 2010 and, in 2015, he obtained a Ph.D. in Mathematics from Universitat Politècnica de València. His research activity focus on Topology and Applied Mathematics. Concretely, in the study of generalized metrics and its applications.

MATEU MORRO is analyst and creative of data models and dashboards to help decision making of our business partners across all the organization, in Data analyst within Commercial Intelligence and Data Analytics department at Hotelbeds Group. He received his B.S. degree from Universitat de les Illes Balears in 2019. His research activity focuses on improving clustering techniques to be able to apply them to model real life problems in sectors such as Tourism, Economy, Business and Medicine.

OSCAR VALERO is Associate Professor at Universitat de les Illes Balears (Departament de Ciències Matemàtiques i Informàtica). He received his M.Sc. in Mathematics from Universitat de València in 2000. Later on, in 2003, he obtained a Ph.D. in Mathematics from Universitat Politècnica de València in the Doctoral Program of Multidisciplinary Mathematics. His research activity focus on Mathematical Models applied to Computer Science, Artificial Intelligence, Engineering, Economics and Medicine.

RECURSIVE ALGORITHMS FOR THE GENERALIZED VANDERMONDE MATRIX DETERMINANTS*

Jerzy S. Respondek
Silesian University of Technology
Faculty of Automatic Control, Electronics and Computer Science, Institute of Computer Science
ul. Akademicka 16, 44-100 Gliwice, Poland
E-mail: Jerzy.Respondek@polsl.pl

KEYWORDS

Numerical Recipes, Numerical Algebra, Linear Algebra, Generalized Vandermonde Matrix, C++

ABSTRACT

In this paper, we derive recursive algorithms for calculating the determinant of the generalized Vandermonde matrix. The main advantage of the recursive algorithms is the fact that the computational complexity of the presented algorithm is better than calculating the determinant by means of classical methods, developed for the general matrices. The results of this article do not require any symbolic calculations and therefore can be performed by a numerical algorithm implemented in a specialized (like Matlab or Mathematica) or general purpose programming language (C/ C++/ Java/ Pascal/ Fortran etc).

INTRODUCTION

The new contributions in this article are the following:

- We derive recursive algorithms for calculating the determinant of the generalized Vandermonde matrix
- The importance of the recursive algorithms becomes clear when we consider practical implementation of the GVM; they are useful each time we add a new interpolation node or a new root of a given differential equation in question.
- The recursive algorithms, which we propose in this work, can allow to avoid the recalculation of the determinant and/or inverse.
- The main advantage of the recursive algorithms is the fact that the computational complexity of the presented algorithm is of the *O(n)* class for the computation of the determinant.

The results of this article do not require any symbolic calculations and therefore can be performed by a numerical algorithm implemented in a high level (like Matlab or Mathematica) or low level programming language (C/ C++/ Java/ Pascal/ Fortran etc).

This work was supported by Statutory Research funds of Institute of Informatics, Silesian University of Technology, Gliwice, Poland (BK/204/RAU2/2019)

PRACTICAL IMPORTANCE OF THE GENERALIZED VANDERMONDE MATRIX

In this article we consider the generalized Vandermonde matrix (GVM) of the form proposed by El-Mikkawy in the article (Mikkawy 2003). For the $n \in Z_+$ real pair wise distinct roots $c_1,...,c_n$ and the real constant coefficient k we define the GVM as follows:

$$V_G^{(k)}(c_1,...,c_n)=\begin{bmatrix} c_1^k & c_1^{k+1} & \cdots & c_1^{k+n-1} \\ c_2^k & c_2^{k+1} & \cdots & c_2^{k+n-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_n^k & c_n^{k+1} & \cdots & c_n^{k+n-1} \end{bmatrix}$$

These matrices arise in a broad range of both theoretical and practical issues. Below we surveyed the issues which require the use of the generalized Vandermonde matrices.

- Linear, ordinary differential equations (ODE): the Jordan canonical form matrix of the ODE in the Frobenius form is a generalized Vandermonde matrix (Gorecki 1993 pp.86-95).
- Control issues: investigating the so-called controllability (Klamka 1991) of the higher order systems leads to the issue of inverting the classic Vandermonde matrix (Respondek 2008) (in the case of distinct zeros of the system characteristic polynomial) and the generalized Vandermonde matrix (Respondek 2008) (for systems with multiple characteristic polynomial zeros). As the examples of the higher order models of the physical objects we can mention Timoshenko's elastic beam equation (Timoshenko 1955) (4th order) and Korteweg-de Vries's equation of waves on shallow water surfaces (Sayed 2004) (3rd, 5th and 7th order).
- Interpolation: apart from the ordinary polynomial interpolation with single nodes, we consider the Hermite interpolation, allowing multiple interpolation nodes. This issue leads to the system of linear equations, with the generalized Vandermonde matrix (Kincaid 2001 pp. 363-373).
- Information coding: the generalized Vandermonde matrix is used in coding and decoding information in the Hermitian code (Lee 2010).
- Optimization of the non-homogeneous differential equations (Gorecki 1968).

ALGORITHMS FOR THE GENERALIZED VANDERMONDE MATRIX DETERMINANT

In this chapter we propose a library of recursive algorithms for the calculation of the generalized Vandermonde matrix determinant. These algorithms solve the following set of practically important, incremental problems:

A) Suppose we have the value of the Vandermonde determinant for a given series of roots $c_1,...,c_{n-1}$. How can we calculate the determinant after inserting another root into an arbitrary position in the root series, without the need to recalculate the whole determinant? This problem corresponds to the situation which frequently emerges in practice, i.e. adding a new node (polynomial interpolation) or increasing the order of the characteristic equation (linear differential equations solving, optimization and control problems).

B) Contrary to the previous one: We have the Vandermonde determinant value for a given roots series $c_1,...,c_n$. We remove an arbitrary root c_q from the series. How to recursively calculate the determinant in this case? The examples of real applications from the previous point also apply here. The proper solution is given in section 3.1.

C) We are searching for the determinant value, when in the given roots series $c_1,...,c_n$ we change the value of an arbitrarily chosen root.

D) We are searching for the determinant value, for the given roots series $c_1,...,c_n$, calculated recursively.

The following theorem is the main tool to construct the above recursive algorithm.

Theorem 1

The following recursive formula is fulfilled for the generalized Vandermonde matrix:

$$\det V_G^{(k)}\left(c_1,...,c_n\right)=\left(-1\right)^{q+1}c_q^k\cdot \det V_G^{(k)}\left(c_1,...,c_{q-1},c_{q+1},...,c_n\right)\cdot \prod_{i=1,i\neq q}^{n}\left(c_i-c_q\right),\quad q=1,..,n$$

Proof of th. 1

Applying the standard determinant linear properties we can obtain:

$$\det V_G^{(k)}\left(c_1,...,c_n\right)= =\det\begin{bmatrix} c_1^k & c_1^{k+1} & \cdots & c_1^{k+n-1}\\ \vdots & \vdots & \ddots & \vdots\\ c_q^k & c_q^{k+1} & \cdots & c_q^{k+n-1}\\ \vdots & \vdots & \ddots & \vdots\\ c_n^k & c_n^{k+1} & \vdots & c_n^{k+n-1}\end{bmatrix}= =\det\begin{bmatrix} c_1^k & c_1^k c_1-c_1^k c_q & \cdots & c_1^k c_1^{n-1}-c_1^k c_1^{n-2}c_q\\ \vdots & \vdots & \ddots & \vdots\\ c_q^k & 0 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ c_n^k & c_n^k c_n-c_n^k c_q & \vdots & c_n^k c_n^{n-1}-c_n^k c_n^{n-2}c_q\end{bmatrix}$$

Next, in compliance with the Laplace's expansion formula applied to the q-th column, we directly have:

$$\det V_G^{(k)}\left(c_1,...,c_n\right)=\left(-1\right)^{q+1}c_q^k= =\det\begin{bmatrix} c_1^k\left(c_1-c_q\right) & c_1^{k+1}\left(c_1-c_q\right) & \cdots & c_1^{k+n-2}\left(c_1-c_q\right)\\ \vdots & \vdots & \ddots & \vdots\\ c_{q-1}^k\left(c_{q-1}-c_q\right) & c_{q-1}^{k+1}\left(c_{q-1}-c_q\right) & \cdots & c_{q-1}^{k+n-2}\left(c_{q-1}-c_q\right)\\ c_{q+1}^k\left(c_{q+1}-c_q\right) & c_{q+1}^{k+1}\left(c_{q+1}-c_q\right) & \cdots & c_{q+1}^{k+n-2}\left(c_{q+1}-c_q\right)\\ \vdots & \vdots & \ddots & \vdots\\ c_n^k\left(c_n-c_q\right) & c_n^{k+1}\left(c_n-c_q\right) & \cdots & c_n^{k+n-2}\left(c_n-c_q\right)\end{bmatrix}= =\left(-1\right)^{q+1}c_q^k\det\begin{bmatrix} c_1^k & c_1^{k+1} & \cdots & c_1^{k+n-2}\\ \vdots & \vdots & \ddots & \vdots\\ c_{q-1}^k & c_{q-1}^{k+1} & \cdots & c_{q-1}^{k+n-2}\\ c_{q+1}^k & c_{q+1}^{k+1} & \cdots & c_{q+1}^{k+n-2}\\ \vdots & \vdots & \ddots & \vdots\\ c_n^k & c_n^{k+1} & \cdots & c_n^{k+n-2}\end{bmatrix}\cdot\prod_{\substack{i=1\\ i\neq q}}^{n}\left(c_i-c_q\right)$$

This concludes the proof.

Q.E.D.

Directly from the Theorem 1 we can obtain the following algorithms for the incremental problems A-D. The detailed implementation of these formulas is straightforward and will be omitted. (2)

Cases A, B: All we need to do is to apply the equality (2).

Case C:

Let us assume that for the given roots series $c_1,...,c_n$ the corresponding determinant value is equal to $\det V_G^{(k)}\left(c_1,...,c_q,...,c_n\right)$. Our objective is to find the value of the determinant $\det V_G^{(k)}\left(c_1,...,c_q+\Delta c_q,...,c_n\right)$. Applying twice the formula (2) we can receive the following expression for the searched determinant:

$$\det V_G^{(k)}\left(c_1,...,c_q+\Delta c_q,...,c_n\right)=\left(\frac{c_q+\Delta c_q}{c_q}\right)^k\cdot$$

$$\cdot\frac{\prod_{i=1,i\neq q}^{n}\left(c_i-c_q-\Delta c_q\right)}{\prod_{i=1,i\neq q}^{n}\left(c_i-c_q\right)}\det V_G^{(k)}\left(c_1,...,c_q,...,c_n\right),\quad q=1,..,n$$

Case D

The proper recursive function expressing the determinant value, for the given roots series $c_1,...,c_n$, has the form:

$$\det V_G^{(k)}\left(c_1,...,c_q\right)=\begin{cases}(-1)^{q+1}c_q^k\cdot\det V_G^{(k)}\left(c_1,...,c_{q-1}\right)\cdot\\ \quad\cdot\prod_{i=1}^{q-1}\left(c_i-c_q\right),\ \text{for } q>1\\ c_1^k,\qquad\qquad\qquad\text{for } q=1\end{cases}$$

Computational complexity of the proposed algorithms

The following facts are worth to be noticed:

- The computational complexity of the presented algorithms A-C is of the *O(n)* class with respect to the number of floating point operations necessary to perform. This enables us to solve efficiently the incremental Vandermonde problems, avoiding the quadratic complexity, typical in the Vandermonde field, e.g. (Eisinberg 2006, Yan 2009)
- The algorithm D is of the $O(n^2)$ class, being, by the linear term, more efficient than the ordinary Gauss elimination method.

Special cases

In this section we give special forms of the algorithms A-D tuned for two special cases of the generalized Vandermonde matrix, i.e. for the equidistant roots as well as the roots equal to the succeeding positive integers.

Generalized Vandermonde matrix with equidistant roots

Let us take into account the GVM with the equidistant roots of the form $c_i=c_1+(i-1)h,\ h\in R$. In this special case formula (2) becomes

$$\det V_G^{(k)}\left(c_1,...,c_n\right)=(q-1)!\,(n-q)!\cdot$$
$$\cdot h^{n-1}c_q^k\ \det V_G^{(k)}\left(c_1,...,c_{q-1},c_{q+1},...,c_n\right),\quad q=1,..,n$$

and the algorithms A-D change respectively to the recursive equation (3).

Generalized Vandermonde matrix with positive integer roots

In article (Mikkawy 2003) it is considered a special case of GVM, that can be obtained from (1) when $c_i=i,\ i=1,...,n$, denoting this matrix by $V_S^{(k)}(n)$:

$$V_S^{(k)}\left(1,...,n\right)=\begin{bmatrix}1 & 1 & \cdots & 1\\ 2^k & 2^{k+1} & \cdots & 2^{k+n-1}\\ \vdots & \vdots & \ddots & \vdots\\ n^k & n^{k+1} & \cdots & n^{k+n-1}\end{bmatrix}$$

For the special $V_S^{(k)}$ formula (3) becomes

$$\det V_S^{(k)}\left(1,...,n\right)=(q-1)!\,(n-q)!\,c_q^k\cdot$$
$$\cdot\det V_S^{(k)}\left(1,...,q-1,q+1,...,n\right),\quad q=1,..,n$$

THE PERFORMANCE TESTS

On the graphs below we practically compared the efficiency of the recursive algorithms with the standard algorithms[1] embedded in Matlab®. On the left we can see the execution time of the standard and recursive algorithms, and on the right the relative performance gain (recursive algorithms vs the standard ones for: inversion and determinant calculation).

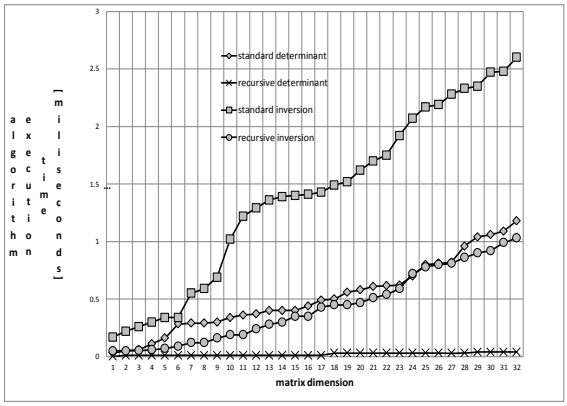


Fig.1 The execution time of the standard and recursive

[1] In a non-recursive form, for matrices with an arbitrary entries - contrary to the algorithms presented in this article, which are developed specially for the GVM.

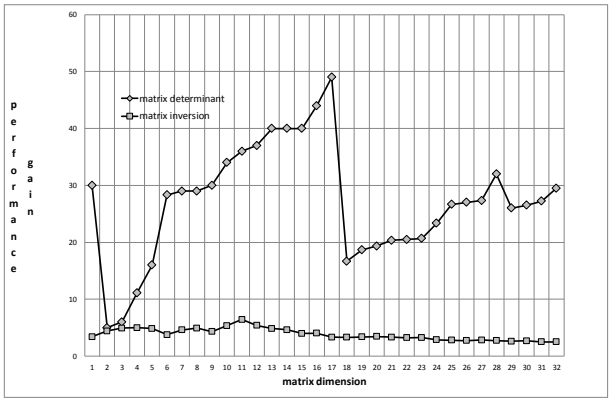


Fig. 2 The relative performance gain algorithms

EXAMPLE

We show practical application of the algorithm from this article on the same numerical example as in the paper (Mikkawy 2003 p. 649), to enable easy comparison of the two opposite algorithms: classical and recursive. Let us consider the generalized Vandermonde matrix $V_G^{(k)}(n)$ with the following parameters:

- general exponent: $k=0.5$
- size: $n=7$
- roots: $c_i=i,\ \ i=1,...,7$

The generalized Vandermonde matrix of such parameters has the following form:

$$V_G^{(0.5)}(7)= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \sqrt{2} & 2\sqrt{2} & 4\sqrt{2} & 8\sqrt{2} & 16\sqrt{2} & 32\sqrt{2} & 64\sqrt{2} \\ \sqrt{3} & 3\sqrt{3} & 9\sqrt{3} & 27\sqrt{3} & 81\sqrt{3} & 243\sqrt{3} & 729\sqrt{3} \\ 2 & 8 & 32 & 128 & 512 & 2048 & 8192 \\ \sqrt{5} & 5\sqrt{5} & 25\sqrt{5} & 125\sqrt{5} & 625\sqrt{5} & 3125\sqrt{5} & 15625\sqrt{5} \\ \sqrt{6} & 6\sqrt{6} & 36\sqrt{6} & 216\sqrt{6} & 1296\sqrt{6} & 7776\sqrt{6} & 46656\sqrt{6} \\ \sqrt{7} & 7\sqrt{7} & 49\sqrt{7} & 343\sqrt{7} & 2401\sqrt{7} & 16807\sqrt{7} & 117649\sqrt{7} \end{bmatrix}$$

The determinant of the matrix $V_G^{(0.5)}(7)$ have the form:

$$\det\left[V_G^{(0.5)}(7)\right]=298598400\sqrt{35}$$

OBJECTIVE

Our objective is to find the determinant of the generalized Vandermonde matrix $V_G^{(0.5)}(8)$, with roots $c_i=i,\ \ i=1,...,8$, in the recursive way.

Recursive determinant calculation

We will calculate the determinant value of the matrix $V_G^{(0.5)}(8)$ by (5) because the GVM in question has consecutive integer roots. In this case the equality (5) leads to the following determinant value:

$$\begin{aligned} &\det V_G^{(0.5)}(8)= \\ &=(8-1)!\,(8-8)!\,\sqrt{8}\,\det V_G^{(0.5)}(7)= \\ &=7!\sqrt{8}\cdot 298598400\sqrt{35}= \\ &=3009871872000\sqrt{70} \end{aligned}$$

Summary of the example

In this example we recursively calculated the determinant of the $V_G^{(0.5)}(8)$ matrix, making use of the known determinant of the $V_G^{(0.5)}(7)$ matrix. It is worth to notice that to perform this there were merely 8 scalar multiplications necessary for the determinant. It confirms high efficiency of the recursive approach.

Summary

In this paper, we derive recursive numerical recipes for calculating the determinant of the generalized Vandermonde matrix. The results presented in this article can be performed automatically by a numerical algorithm in any programming language. The computational complexity of the presented algorithms is better than the ordinary GVM determinant methods.

The following can be seen as the desired future research directions:

- Construction of the parallel algorithm for the generalized Vandermonde matrices.
- Adaptation of the algorithms to the vector-oriented hardware units, like Intel AVX and SSE.
- Combination of both.

REFERENCES

R. Bellman. 1960. “Introduction to matrix analysis”, Mcgraw-Hill Book Company, New York.

A. Eisinberg, G. Fedele. 2006. “On the inversion of the Vandermonde matrix”, Appl. Math. and Comp. 174, 1384-1397.

S. M. El-Sayed, D. Kaya. 2004. “An application of the ADM to seven-order Sawada-Kotara equations”, Appl. Math. and Comp. 157, 93-101.

H. Gorecki. 1968. “On switching instants in minimum-time control problem. One-dimensional case n-tuple eigenvalue”, Bull. de L’Acad. Pol. Des. Sci. 16, 23-30.

H. Gorecki. 1993. “Optimization of the dynamical systems”, PWN, Warsaw.

S. Hou, E. Hou. 2007. “Recursive computation of inverses of confluent Vandermonde matrices”, The Electron. J. of Math. and Technol. 1, 12-26.

S. Hou, W. Pang. 2002. “Inversion of confluent Vandermonde matrices”, Comput. and Math. with Appl. 43, 1539-1547.

D. R. Kincaid, E.W. Cheney. 2001. “Numerical Analysis: Mathematics of Scientific Computing”, Brooks Cole, third ed., California.

J. Klamka. 1991. “Controllability of dynamical systems”, Kluwer Academic Publishers, Dordrecht.

K. Lee, M. E. O'Sullivan. 2010. “Algebraic soft-decision decoding of Hermitian codes”, IEEE Trans. on Inf. Theory 56 (6), 2587-2600.

M. E. A. El-Mikkawy. 2003. “Explicit inverse of a generalized Vandermonde matrix”, Appl. Math. and Comput. 146, 643-651.

W. H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery. 1992. “Numerical Recipes in C”, second ed., Cambridge University Press, Cambridge.

J. Respondek. 2008. “Approximate controllability of infinite dimensional systems of the n-th order”, Int. J. Appl. Math. Comput. Sci. 18 (2), 199-212.

J. Respondek. 2008. “Approximate controllability of the n-th order infinite dimensional systems with controls delayed by the control devices”, Int. J. Syst. Sci. 39 (8), 765-782.

J. Respondek. 2011. “On the confluent Vandermonde matrix calculation algorithm”, Appl. Math. Lett. 24, 103-106.

J. Respondek. 2011. “Numerical recipes for the high efficient inverse of the confluent Vandermonde matrices”, Appl. Math. Comp. 218 (5), 2044-2054.

S. Timoshenko. 1955. “Vibration Problems in Engineering”, D. Van Nostrand Company, third Edition, London.

S. Yan, A. Yang. 2009. “Explicit algorithm to the inverse of Vandermonde matrix”, Int. Conf. Test and Measurements, 176-179.

AI-BASED DATA ANALYSIS

THE EVOLUTION OF ARTIFICIAL INTELLIGENCE TOWARDS AUTONOMOUS SYSTEMS WITH PERSONALITY SIMULATION

Joël Colloc
Normandie Univ, UNIHAVRE, UNICAEN, UNIROUEN, CNRS IDEES
email: joel.colloc@univ-lehavre.fr

KEYWORDS
Autonomous systems, Ethics, Robot personality, Emotion modeling, Responsibility, Forensics, Legality

Abstract

Evolutions of artificial intelligence towards autonomous systems with personality-simulating devices, artificial emotions and languages will increasingly interact with human users and users and become partners in their everyday lives. This paper will examine the consequences of these interactions both at the individual level and on the place that these autonomous systems will take in the society. In February 2017, the European Parliament approved a report presented by an European deputy. This report analyses the legal consequences of the growing presence of robots in European daily life. A key measure proposed in the text is to give legal personality to robots. According to industry partners and stakeholders, this evolution of the law is made necessary by the precautionary principle and by the challenges of competitive and growth advantages. Are these systems useful to humans and acceptable to them ? The presentation presents the main technological advances and feasibility of an artificial and emotional personality for robots, the ethical aspects of robots' interactions with humans, in particular the medico-legal notion of responsibility to individuals and the societal consequences of robotic deployment in society.

Introduction

This article will focus on the evolution from memetic (MAI) to autonomous artificial intelligence (AAI) with the important question of machine that can be able to modify itself, to set up by itself its objectives and to decide alone how to accomplish them ? Moreover, in the future some researchers trust that a machine could develop an artificial way of thinking making a robot to become an artifical person with legal personnality, responsible of the dammages caused by their acts to people and goods. A lot of stakeholders and researchers develop AI systems in medicine. Thus, we will take the healthcare domain to illustrate the evolution of AI and their consequences for patients and caregivers. The article is structured in five sections: the first section describes the evolution from memetic AI to autonomous AI, the second section concerns the feasibility of autonomous systems and what features must be implemented to simulate artificial thinking capabilities. The third section presents the motivation to design and to develop autonomous robotics. The part four describes the impact on citizen life and particularly the social, legal consequences and a toolbox for digital ethics to take appropriate decisions. The next section try to forsee consequences of autonomous AI and robots endowed with consciouness and transhumanism.

EVOLUTION OF ARTIFICIAL INTELLIGENCE: FROM MEMETIC TO AUTONOMOUS SYSTEMS

The AI is located at the crossroads of four families of disciplines: cognitive science that studies the nature of human intelligence: psychology, Linguistics, sociology, philosophy) neuroscience dealing with the operation of the brain, perception, emotions. Biology that describes the phenomena of living including genetics and social insects (Entomology), the behavior of the animals (ethology). Artificial learning (Machine learning) that uses mathematics and computer models to model learning, pattern recognition, knowledge modeling. Jean-Louis Le Moigne describes the evolution of information systems and a classification inspired from K. Boulding (1956). Like in epigenesis, the following level adds features to the previous one. By conciseness, only the levels 6 to 9 are presented Figure 1.

Information hierarchy

On level 6, the system becomes able to memorize his decision O.I.D (Operating system (OS), Information system (IS), System of Decision (DS)). On level 7, the system coordinates numerous decisions of actions at all the time t, concerning its internal activity, regulation and the external information from and to its environment (SCS). On level 8, the system is endowed with a subsystem of imagination and design (SID). On level 9, the system is able to decide on its decision and to determine

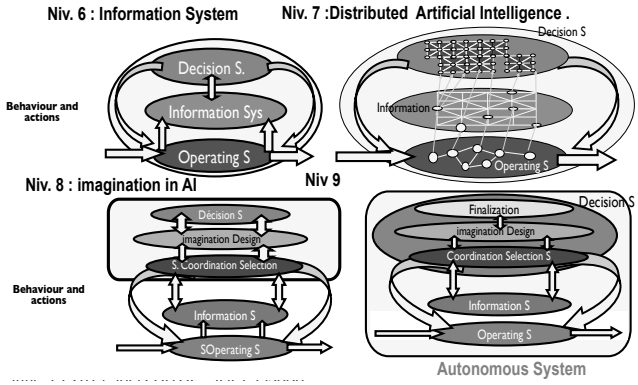


Figure 1: Complex system classificafion J-L. Le Moigne

the positive and negative aspects of its actions. This finalization of a complex system is close to the human thought (FS) which confers it an autonomy of decision allowing it to set its own goals. Autonomous multi-agents systems (AMAS) belongs to this kind Le Moigne (1990).

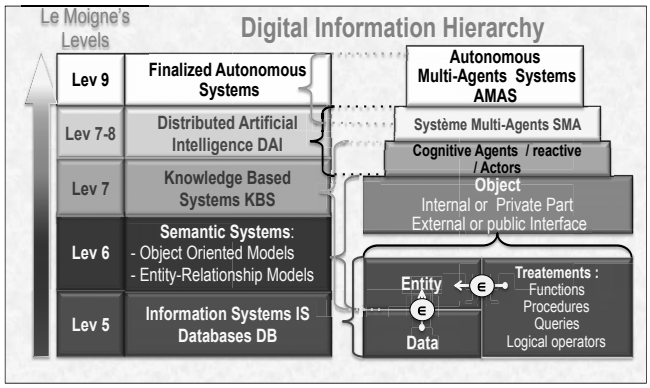


Figure 2: Digital Information Hierarchy

Information hierarchy

Information relies on the triptych: data, information and knowledge Abiteboul (2012). But this classification must be completed with two important concepts: digital trace Galinon-Mélénec (2011) and metaknowledge Pitrat (1990) Figure 2. AI is not just data but involves a hierarchy of information sources: On figure 3 from bottom to top: AI is not just data but involves a hierarchy of information sources: from bottom to top: -the digital trace is left intentionally or not on the net by the user: The clicks provide information on the preferences of users. They are time-stamped by the website. Data are characteristics or properties of a user and the value belongs to an integrity set. Knowledge are Guidlines, heuristics are efficient methods to solve a problem in the same type of situation, they are indepententes from individuals.Meta-knowledge determines what type of knowledge is necessary and how to use it to solve a specific type of problem. For example how to make a diagnosis. Obviously, meta-knowledge is also independent from individuals. Information is the inter-

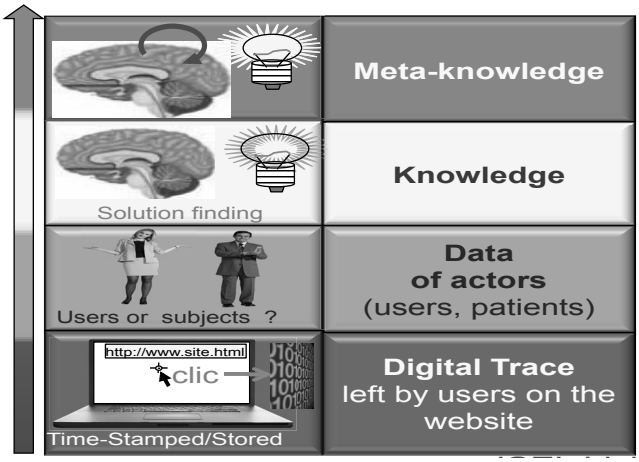


Figure 3: Hierarchy of Information sources from AI

pretation of one or more data associated with an object or actor (for example a patient in an hospital) whose treatment creates meaning in a context with arithmetic $+, -, *, /$,logical {and,or,xor,not} or comparison operators $\{=, <, \leq, >, \geq, \neq\}$. The DateOfBirth attribute is used to compute the age of the patient with the function $Age()$. Th attriute Glycemia to assess the risk of diabetes by computing the $DiabetesRisk()$ function when it exceeds the threshold of $1.26 gram/liter$ during two successive measurements.

Knowledge: In short, knowledge is the result of epidemiological studies in the population people and founded an aid to the diagnosis, prognosis, treatment. Databases index data collected and analyzed for each identified patient (medical record). The patient leaves traces of use on the Internet and social networks that allow a secondary re-identification and consolidate more data. For Alain Mille, "observation is a cognitive process to distinguish the imprint left in the environment (in the proper sense) as trace of something that may make sense... A paradox of security appears very quickly: access to the resources available in the computerized environment (not only the informatics environment so) very often requires procedures for identification of actors... Associated imprints are therefore very often specifically associated with the activity of its users and the observed traces reveal private behaviors."...Mille (2013). The management of the property of computer traces is a nontrivial issue. Increasingly, operators or companies who offer software or licensing systems, condition their utilization to the abandonment by the user of the property of its personal data and tacitly grant a right of communication to third parties for commercial purposes.

Memetic AI, Knowledge base and decision support systems

AI today is a catch-all concept. However, it includes two main tracks: the Mimetic AI (the initial way) and autonomous AI (the new way). There are three steps in the Mimetic AI, the first step is the transfer of knowledge from experts to build a knowledge base in a given area, the second step is the designer of the knowledge base provides a support system to the decision Figure 3. The third step is the health care user (for example a medical student) dialogues with the system to solve clinical cases. In medicine, many systems have been developed which render services in aid to diagnosis and prescription for example assistance to users in areas or territories where few physicians are available. Convincing medical simulators are offered for students to place them in a situation of care without risk to the patient. For example SIAMED offers a prescription from the diagnosis of infectious disease and its location in five successive stages which exploit the knowledge base of the antibiotics and clinical data of the patient's medical record Figure 5.The system implements the following clinical steps : the indications binded to the bacteries that could cause the disease and its localisation, the contraindications related to the State of the patient, drug-drug interactions with current treatments and takes into account the risks of side effects for each antibiotic. From the selected molecules, it offers the specialties of drug with a dosage and route of administration (Colloc 1985). Further works extend the capabilities of the system to infectious disease diagnosis (Shen et al. 2018). Two inseparable aspects of medicine coexist, complement themselves: firstly, the clinical medicine in the service of the patient, each individual in the relationship (caregiver-patient) care with doctors on the left part of Figure 6. Secondly, the public health at the service of all of the population. The public health focuses on risk factors, to epidemics, to bad behaviour based on statistics in the population. It issues recommendations, hygiene measures and means of prevention on the right part of Figure 6.. The objectives of public health and clinical medicine are sometimes contradictory. Care cannot be imposed on patients. In short, medical knowledge are the result of epidemiological studies in the population people and founded an aid to the diagnosis, prognosis, treatment. Knowledge comes also from medical publications that precisely describe the deseases and compare the treatment efficiency from the clinical experience of physician mainly in hospitals. Data are indexed in databases and are collected and analyzed for each identified patient (medical record). During the same time, the patient leaves traces of use on the Internet and social networks that allow a secondary re-identification and consolidate more data concerning the health and personal data of each person.For example, he will consult a medical website or a medical forum to understand the signification of clinical signs or to get information concerning the treatment of his/her disease Figure 7.

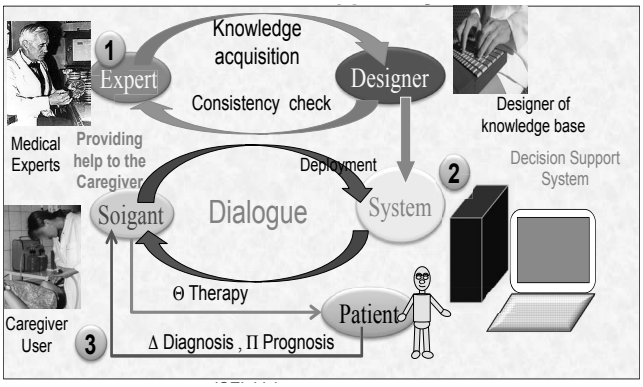


Figure 4: Memetic IA Decision support system

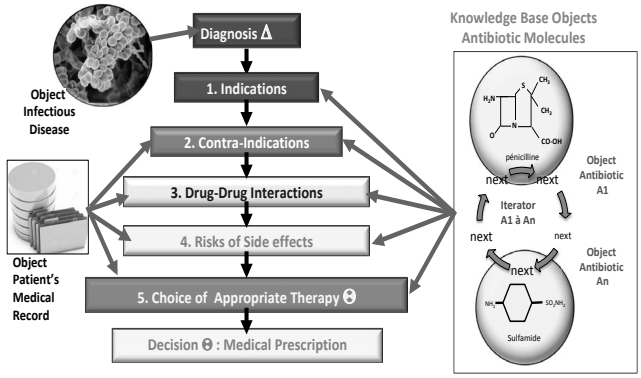


Figure 5: SIAMED : Antibiotherapy decision support system Colloc (1985)

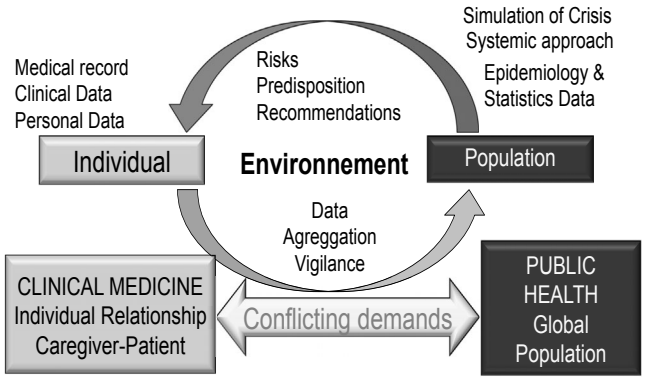


Figure 6: Clinical Medicine and Public Health

Seven W to profile the Internet user

According to the method of Marcus Fabius Quintilianus (35-96 apr J.-C): who is doing what, why, how, with what, where and when. How Big Data turns digital traces of a hypothetical user (or patient) into data? The

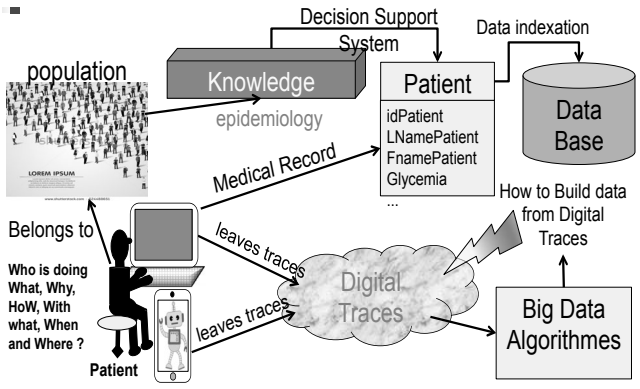


Figure 7: Reidentification of the Web surfer

user profiling is determined by its repeated uses: Who : User identifier : IP Adress or email What : What website is visited by the surfer, Why: The traces leaved by the user (clics and cookies) allow to know the motivations of the Internet Surfer for visiting the website. How: Clicks and keyword of research used by the Internet Surfer determines his/her interests in the visited website With What: What Computer, Tablet, Pad or Smartphone is used. Where: The Website gets the geolocation of the machine When: The user action (query or click) is logged and timestamped. However, the actual user is uncertain, because this is the machine that is identified and several users (e.g. of the same family) may use it in turn.

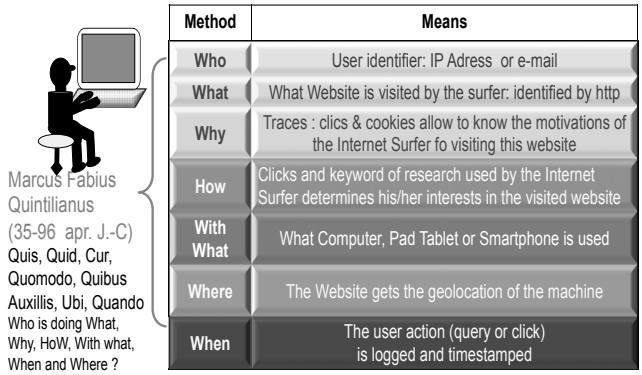


Figure 8: The Web surfer profiling

Deep learning

The deep-learning aims to classify a large number of images or profiles of people from the net, in an autonomous way (Big-Data). He uses an artificial neural network with a large number of layers. In spite of his surprising performances, the main inconvenient of deep learning is that it doesn't implement meta-connaissance to explain the basis of the knowledge obtained and how and when to use this knowledge. John Searle's Chinese room: a man closed in the chamber is doing the matches between input questions with the output answers both written in chinese with the help of a manual of rules. The man could not pretend that he is able to master the chinese. In fact, to recognize does not mean to understand.

HOW TO BUILD AUTONOMOUS SYSTEMS?

The recent progress of the cognitive sciences, neurosciences, computing and robotics boosts the project of autonomous artificial being. The deep learning methods allow to enhance the artificial perception (sight, hearing, touch, smell, proprioceptive sensitivity). However cognition especially reasoning and thinking remain a big challenge. The feasibility of autonomous artificial thinking systems needs to compare the way the human beings acquire their information and develops the thought with the current capacities of the autonomous information systems. We proposed a model based on four hierarchies: the Le Moigne's hierarchy of information systems supplies indicators of complexity and autonomy; the cognitive hierarchy describes the sub-symbolic acquisition and the emergence of our personal experience, our knowledge, whereas the linguistic hierarchy builds the speech describing the knowledge acquired in terms of concrete and abstract objects on the environment and on oneself Figure 14. The digital information hierarchy relies on the necessary concepts, models, methods and tools to build autonomous information systems (Colloc 2016). The production of autonomous systems is rather different than memetic AI systems and the role of the end-user is totally changed Figure 9. An investor sells

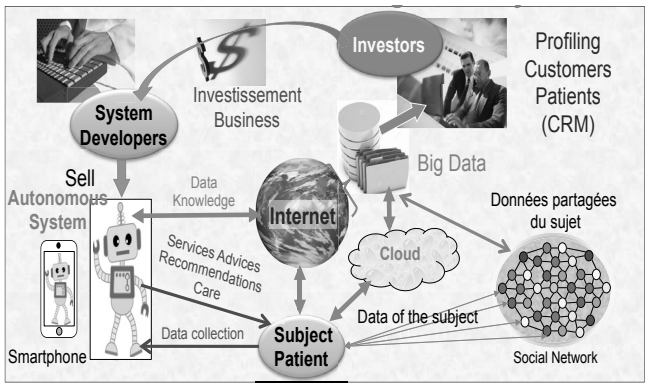


Figure 9: Autonomous AI context

robots companions or smartphone autonomous virtual assistants which provides advice, service, recommendations (in the future some cares) to subjects directly through a dialogue without the help of a caregiver. The autonomic system collects data of the subject and is able to get access and query data and knowledge on the Internet (Cloud, social networks, Big Data). He makes decisions in complex situations for the well-being of the subject. The collection of data allows to better understand the needs and expectations of the subject and

how to answer or even to prevent his needs (may be too much).

The EPICE model

EPICE is a french acronym for (Emotions, Personnalité, Interactions, Connaissances, Expériences) is a model that takes into account (Emotion, Personality, Interaction, Knowledge, Experience) and implements them in a decision support systems of medical ethics Colloc and Léry (2008). The implementation of the layers of the

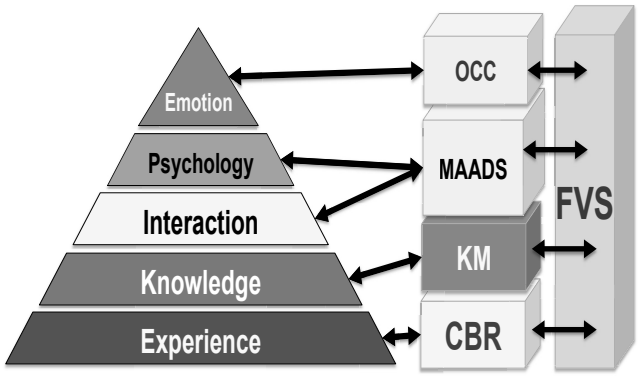


Figure 10: The EPICE Model

model is done with the fuzzy vectoriel space (FVS) Figure 10 which is described in our previous work Colloc and Summons (2017).

The emotion layer E

The emotional layer E is composed of attributes that express balance and define the agent's emotional states. It is based on a the Ortony Clore and Collins (OCC model) proposed in Ortony et al. (1988) and the Roseman's model Roseman (1996). The layer E model defines a set of 11 emotional parameters that represent both positive and negative intensities Figure 11. Each emotion can change rapidly over time. The emotional parameters of the OCC model are represented by fuzzy functions and combined in a fuzzy vectorial space (Colloc 2017)

The personality layer P

The personality layer P describes the psychological and behavioural structure of an actor. These are generic prototypes describing typical psychological structures. These prototypes are specialized and thus adapted to the different actors in order to model the behaviours actually observed. They correspond to decision-making patterns that are likely to evolve over an individual's lifetime. The personality layer is defined by a library of automatons modeling general behaviours and strategies. These automatons define metaknowledge in psychology.

		+	-
Consequences of events	For others	Happy for	Resentment
		Gloating	Pity
	For self	Hope	Fear
		Joy	Distress
Actions of Agents	Self Agent	Pride	Shame
		Gratification	Remorse
		Gratitude	Anger
	Other Agent	Admiration	Reproach
		Gratification	Remorse
		Gratitude	Anger
Aspects of Objects		Love	Hate

Figure 11: The OCC parameters

They model the decision-making modes corresponding to the evolution of an actor's emotional states. The personality layer is in close relation to the emotional layer which allows to represent sympathy, antipathy, aversion, self-esteem. The shape of automatons and the functions of transitions allow to represent particular psychological behaviors. For example, an obsessive actor will be represented by an automaton having a number of feedback loops due to his checking behavior before getting satisfaction.

The interaction layer I

The interaction layer I describes the relationships and interactions that an actor has with other actors in the system. The complexity is linked to the diversity of the actors involved. Each actor can maintain relations with both legal persons and natural persons. The question that arises is what is the place of the various actors within a system of decision-making support in ethics in the context of humanitarian health. Catherine Fuchs in her book revisits the hypothesis formulated by Sapir and Whorf of linguistic relativism Fuchs (2003; 2004). For Edward Sapir, the language reflects a community's vision and interpretationSapir (1930). His work is based on the classification of Native American languages.For Whorf there is no objective or universal reality, but only representations of it that would be determined by language. Catherine Fuchs explains that each language constructs a different "world view" because each language community selects distinct isolates of experiences and gives them shared meaning. Regarding the question of variants and invariants (one-sided): "the diversity of representations constructed through languages is a central question for linguistics: the theory of articulation concerning variations and invariants." We think that the triptych Language, Thought and Cognition is in action here: according to Jakobson, thought differs according to language. For Pinker and Fodor: there is no thought without language? Is there a thought and cognition without language? Do we think the same way

as the Japanese or the Lapps? Louis Léry clearly showed that he was not. The culture of a people is organized around verbal and non-verbal modes of communication, beliefs, customLéry and Colloc (2008) . Figure 3 describes all the areas and constraints that will be involved in the decision. It also shows the hierarchy of decision-making criteria: law, ethics, charters, customs, beliefs and religion. The brain is like a plant that grows in an environment: Our knowledge is linked to our culture and our environment. Can it be transferred to other peoples of different cultures? It is doubtless possible, if one spends long years in a country, to acquire the customs and customs and to forget, unlearn one's previous culture, which will gradually fade away. This is just a hunch based on observation of a few cases. Our person can be fully and develop only in interrelation with others, hence the importance of the ≪relational flower≫. How to model interactions between legal and physical persons, groups of persons? A n-m relationship is necessary in order to represent all the interactions that may exist between many actors. The graph complexity increases dramatically ad $(na.nr)^2$ where na is the number of actors and nr is the number of modeled relationships between them. Models from the work of linguists such as Quillian on semantics have been reused and adapted by computer scientists Quillian (1968) J.F. Le Ny shows the interest of semantic networks Ny (2005). The ontology models are intended for the axiomatisation of a field of knowledge Shen et al. (2015).

The Knowledge layer K (Connaissance C)

The knowledge layer defines the sources of knowledge in which an actor has access because of his roles and of his skills. We define five categories of knowledge in a independent way chosen models of representation and domains of knowledge. Factual knowledge: it is about data describing an object of the world real and generally admitted by all. The observed facts are confidentially connected to the truth and classified according to their degree of certainty and precision. It is about a statistical argument: for example: the majority of the people have a similar perception of a characteristic of an object, for example its color is red. Knowledge heuristics: if a situation S is then observed we have knowledge which are relevant and valid in this situation S. It can be properties of objects or a usually applicable method successfully in this situation. The causes of the validity of a knowledge heuristics are not always available. Procedural knowledge: how acts on the world: know the chains of tasks to be made to reach an expected result. It is about all the procedures or the courses to follow expressed by a list of tasks to be realized to be effective in a given situation. Dynamic or behavioral knowledge: they concern the spontaneous variation of the facts, or a behaviour in time which is usually observed (for example, the earth rotates around its axis once every 24 hours). They are useful for the simulation of natural phenomena. It is about the perception of the various states spontaneously taken by objects during given period, of their interactions. The behavioral knowledge have a major importance in sociology, in economy, in botany, in medicine and in physics where we observe the behavior of a system. Deep learning is the new type of knowledge that can be now implemented in systems . System knowledge modelling relies on operators detecting the similarity of knowledge objects. Other important aspects are time modelling and the reasoning operators: deduction, induction, abduction, sumsumption and analogy that allow to combine, compare knowledge objects. Deduction is based on the modus ponens or modus tollens.Induction tries to propagate a property observed in one object to all objects that belong to the same class and allow to split the class in two subclasses the one with the property and the other without the property. On the contrary, abduction tries to refute a property usually oberved in the objects of a class. Sumsumption tries to generalise properties to more general concepts: Think of the individual under the general (an individual under a species, a species under a genus); consider a fact as understood under a law. General sumsumption could be applied to implement induction as proposed by Buntine in 1988 Buntine (1988). Analogy transposes the relationships and properties of objects from one universe to another one provided that these universes and object classes are sufficiently similar.The similarity is computed by a distance.

The Experience layer E

The experience allows to enhance the knowledge in a specific area of the science. Case-based reasoning (CBR) implements a kind of analogy. CBR is a model of experience that allows to index and store cases in an object-oriented database and then, in next consultation, to use a distance to retrieve in the database the more similar cases to the new case in order to apply the most appropriate methods of the previously stored cases to solve the new problem Kolodner (1983)**?**Kolodner (1993)(Aamodt and Plaza 1994). CBR involves semantic distances developed by different approaches: algorithms of structural similarity Boulanger and Colloc (1992); statistical learning as proposed by Zighed D.A. and G. (1992); digital approaches from neural networks and fuzzy logic. The distances are using to implements the different stages of the CBR cycle as depicted on Figure 12 Colloc et al. (2007) The researches about semantic distances tend to combine symbolic and numerical aspects Shen et al. (2015).

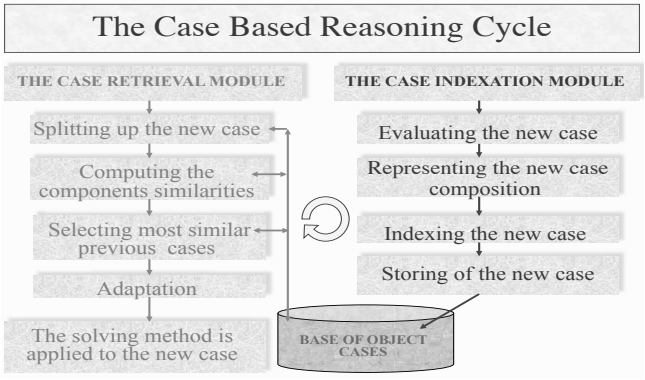


Figure 12: A Case Based Reasoning Cycle

MOTIVATIONS FOR AUTONOMOUS ROBOTICS

The advantages of robotics

The advantages are the ones of the cybernetics or robotics which aims at the improvement of the human condition by unloading people of hard work and to assist them in the daily life as diligent and tireless companions in any sorts of activities of the life: education, supervision, safety of the elderly, the patients, the children. They are endowed with emotion and have a behavior which is going to look like more and more that of the human being, what makes the interaction with them much more pleasant. The Conscious Autonomous Multi-agent Systems (CAMAS) has an internal memory to store facts of consciousness, but especially, they reach the global memory of the humanity from now on stored in Internet: not only in the knowledge but also in the data concerning each of us establishing our profile of way of life. The main question is adressed in the following subsection.

What is thinking ?

For the linguists as Pinker and his pupil Fodor, there is no thought without language, that is without words indicating the objects of the world. Jerry Fodor proposes an internal language named the "mentalese" which gives to each a reflexive thought, the capacity to speak to to oneself and so to repeat for one of the information useful for his/her life. This reflexive thought is strongly bound to the consciousness of oneself and his/her body, to exist, to be an alive entity of the world with his history, souvenirs, feelings and projects. Even there, the self-awareness: " The I " appears late in the life of the individual, during the childhood and even the adolescence. The capacity to think is previous to the implementation of the language. When we speak to oneself, the internal verbalization stimulates the same intellectual cortical zones as during the expression by the word with request of the driving ways, but the motricity of apophyses aryténoïdes which tighten the vocal cords would be inhibited by the brain preventing the emission of sounds of the voice (Jacquet-Andrieu and Colloc 2014). Thus there it would be no important difference to speak and speak to himself (in the sense of the mentalais of Fodor), in both cases the areas of the language are requested Figure 13. We speak to ourself to strengthen our capacity of analysis and resolution of the problems but this activity is not necessary for the thought and do not even maybe fatal in a perception more lit of the world. Is to speak to oneself constitutes the only way of thinking?

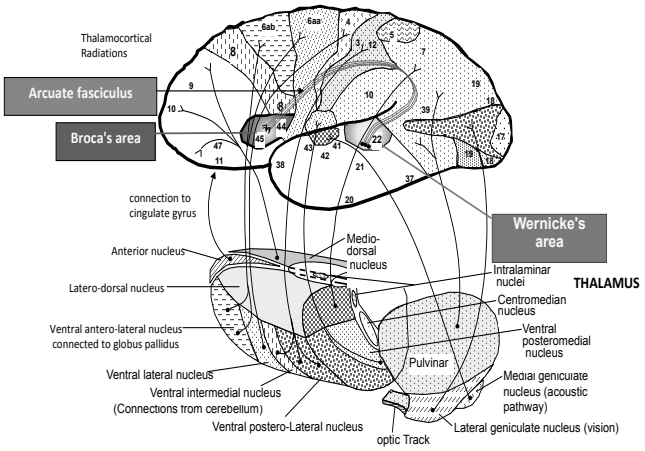


Figure 13: The brain, language area and thalamic connexions

Autonomous systems: Could computer think ?

Computers could think in a very different way than human, as a plane flights differently but faster than a bird Pitrat (1993). The Zen is a total holism, the world cannot be absolutely divided into parts. The dilemma is that for every object of the world, according to master Zen Mummon: "we cannot express him with words and we cannot express him without the words". According to the Buddhism : to trust the words to reach the truth is equivalent of trusting an always incomplete formal system **?**. For Jiddu Krishnamurti, our consciousness is common to all the humanity: All the human beings think that contribute to build it. He considers the individualism, the ego as an obstacle to understand the consciousness with rare moments of clarity ("insight") Krishnamurti (2005). *"The thought is a movement in the time and the space. The thought is memory, memory of the past things. The thought is the activity of the knowledge, the knowledge which was gathered through millions of years and stored in the form of memory in the brain."*Krishnamurti (2005). There are two forms of thinking: -The first is a reaction of the memory which contains the knowledge, the result of the experience from the beginning of humanity (phylogenesis)

and since our birth (epigenesis) in a loop: experience knowledge memory thought action and so on, necessarily limited by the time, it is used every day, rational, individualistic, power-hungry and of progress submissive to the knowledge which accumulates, in the words which divide and this division is responsible for all the suffering, for all the troubles of the world.This first type of thinking relies on the symbolic treatment of the information and the linguistic hierarchy upper right part of Figure 14 with linguistic operators and able to implement the mentalese as described by Jerry Fodor. -The alternative thinking: *the "active-attention" that occurs in rare occasions, we pay simply our attention to the world, without interpreting it, without naming anything, virgin of any prejudice, knowledge and especially spontaneously, by living this moment without thinking of it and without the will. For example: the direct perception of a wonderful landscape of mountain one morning with all our senses uses our complete attention where we forget ourselves and banish the use of the words"* Krishnamurti (2005). Such a full perception requests only

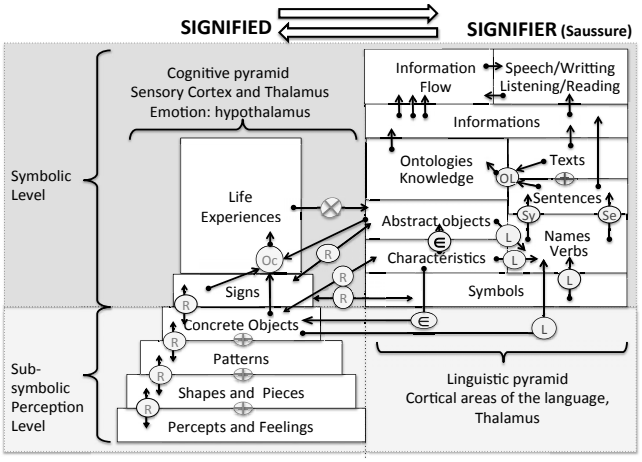


Figure 14: The cognitive pyramid

the left lower part: the sub-symbolic cognitive pyramid and not the linguistic pyramid Figure 14. The first way of thinking is a fatal vicious circle implemented soon in computers that will think better than human beings! And there nothing prevents that a computer invents a new religion at the origin of new sufferings for the humanity. The language during the evolution would have made lose to the man his spontaneity in the immediate perception of the world such as it is. The digital technology, which strengthens the symbolic nature of our relation to the world, is doubtless going to establish the peak of our ignorance. The realization of CAMAS able to think is now possible. Alain Cardon proposes a psychic system that relies on knowledge in psychology and capable of generating flows of thoughts that take place in the temporality under the shape of organized groups of processes to build the artificial psychic system and its interactions. This conscious autonomous multi-agent system (CAMAS) is able to choose its objectives and set its goals to achieve them (level 9 of the Le Moigne's classification of complex systems. Alain Cardon recently published a method to develop systems with psychological states and with the symbolic way of thinking inspired from Sigmund FreudCardon (2016) Cardon (2018). The CAMAS can be reproduced in very large numbers and coordinate each other to reach complex goals fixed in common. The CAMAS is thus going to overtake the human being in his capacity to be thought and to act in the world. Furthermore, it is endowed with mechanical, robotics architecture allowing it to move and to act in the world with more power and efficiency than human beings. It can be endowed with artificial senses like ordinary sight but also infrared vision, large spectral hearing, nano-cameras amplifying its perception far beyond our possibilities. The main advantage would be to be able to improve the capacities of investigation of our world including in hostile circles as the conquest of space by taking advantage of the concept of telepresence without risk Hofstadter and Dennett (1981). These systems have access to Internet and will quickly become more powerful than the human beings with disturbing consequences for the future of the humanity. We have shown that computers can think in the first way only. The main characteristic of humanity is to still be able to think also in the alternative way but how long before losing our essential faculty ?

IMPACT ON CITIZEN LIFE

Medicolegal aspects

Civil responsibility, which is based in France on article 1240 of the new civil code (ex 1382) is different for the mimetic AI and autonomous AI. For example in medicine, mimetic AI tools like a surgery robot is used under the responsibility of the surgeon who uses it. Consequently, the surgeon is responsible for any possible damage. In the context of an autonomous systems that makes alone a care act (example: drug delivery system). If this act causes injury to the patient. The company that produced the Robot is responsible for the damage caused by the system if and only if the link between the act and the damage is proved Figure 15.

Ethics of artificial intelligence

There are two level of ethics: the first is the population and public health level where for example in France, the National Consultative Ethics Committee (CCNE) offers general recommendations useful to the population to address ethical decisions related to health and listening to citizens. Digital clinical ethics proposes a method of ethical decision-making about each patient the relationship of caregiver-patient care at the hospital or the practitioner's office. They try to foresee the

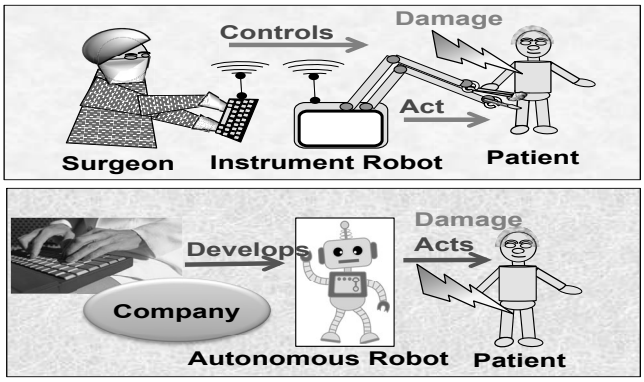


Figure 15: The civil liability of Robots'acts

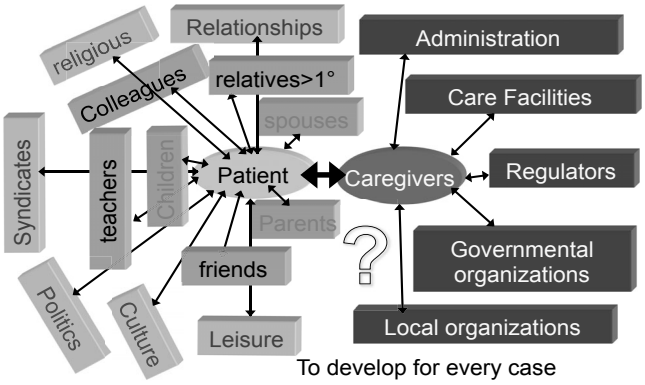


Figure 16: The relational flower of a patient

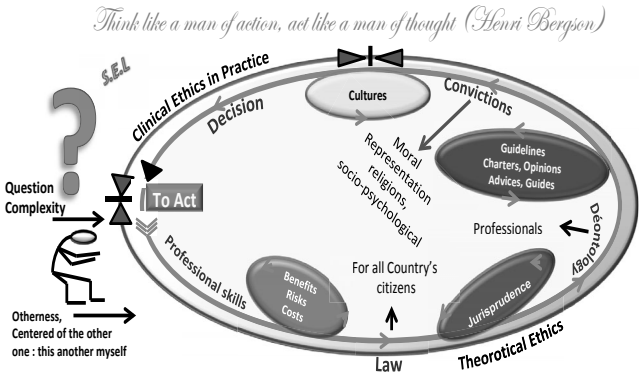


Figure 17: The ethics decision toolbox

impact of digital on the relationship of care and outcomes (benefits and risks) on the life of patients. The second is the clinical level that considers each individual in his environement: Ethics begins where the ready-to-wear of morals is no longer sufficient. Ethics act is first a response to a situation limits and complex from the latin "respondere" with the double meaning: to warrant the acts i am doing to care this person, and to bring an answer to the problem; which is responsibility Léry and Colloc (2008). A Toolbox has been developed by Nicole and louis Léry with whom we formalized a metamodel of ethical decision published in 2008 **?**.The method is pragmatic and is based on the Bergson's motto: thinking as a man of Action, act like a man of thought. Initially a question concerning the patient (this other myself). The question is necessarily intimate, unique, complex, because it concerns that person, his/her history. The question solicits his entourage (relational flower Figure 16) and need a reflection, then a decision that must lead to an action. This path leads us from the theoretical ethics (such as proposed by some Ethics Advisory Committees CCNE in France) to practical clinical ethics by successively visiting the following arguments: technical skills: what can be done with the professional skills in the medical art, knowledge, know-how, then, the risks/benefits/cost equation for assessing the achievable potential solutions. Then there are legal targets: law, jurisprudence and ethics for professionals, the recommendations and charters, moral, religious, beliefs the sociopsychological landmarks, cultures. All of the families, friends, caregivers who coexist the person participate in ethical reflection Figure 17. We are currently working to apply this ethics toolbox to digital ethics and escpecially to autonomous systems.

Societal consequences

The recording information on the populations is former. Totalitarian regimes use it and have now powerful means to profile anybody in very precise way and with facial recognition they control the activity of whole population. The constitution of personal databases under the guise of public health needs have shocked the doctors: for example : the System GAMIN who organized the priori discrimination of socially disadvantaged children (french law of 15 July 1970). This law required that physicians declare disabled children, which constituted a priori discrimination of these children Colloc and Lery (1989),Vitalis (1981). We have seen that in technical terms, the concept of local administered and protected database whose purpose is known, declared and controlled in France by the *Commission Nationale Informatique et liberté* CNIL becomes obsolete due to transborder flows of personal data. This institution has therefore adapted and the french law "informatique et liberté" and has been enriched by the european General Data Protection Regulation (GDPR). But its means remain insufficient and especially its scope is limited to the national or the european perimeter while the Internet is a global network with different rules and laws. Outside european borders the controls and prosecutions remain difficult and most of the time impossible.

Social risks of disclosure of medical data

The disclosure of patient health data may include a wide range of negative consequences. Firstly it can lead to a

"loss of luck" related to the disclosure of disease (HIV Infection, Hepatitis, Cancer, Addictions, state of depression) to a third party (insurer, employer, lender) who will refuse to grant a loan, employment or will increase the insurance fee: because he know the risk inherent to the disease. The exploitation of information against the person in a dispute (divorce, inheritance, childcare,...) can cause unpleasant results. The attacks on the reputation, brand image, ridicule, harassment related to the knowledge of a disease or an impairment... can cause very serious consequences with an exclusion from mainstream society, to be pushed aside in the company, job loss, rejection, marginalization, depression that can lead to suicide.

Societal risks of autonomous systems

The main societal risk is that the CAMAS is put in the service of the ambitions and in the service of the power of an oligarchy of people who can acquire these very expensive sophisticated AI machines and on behalf of economic necessities, to impose their wills, their power on the rest of the humanity with the implementation of a totalitarianism based on the superhuman capacities (physical strength, speed of calculation, strategy, artificial thought) of such machines Figure 9. The study from the ethical point of vue compares the arguments supplied by western and oriental philosophers on the nature of the thought and its role in the evolution of the humanity. The thought occurs at the sub-symbolic level in the cognitive pyramid. According to the Buddhist philosophy, but also the western logicians, giving a name to things is to classify and to classify it is to divide as defined by Lewis Carroll Carroll (1896). This division is source of the conflicts and prevents us from seeing the world directly. We are conditioned in an individualistic thought organized around the "me". The realization of systems conscious autonomous multi-agents endowed with the capacity to think is now possible. These systems will quickly become more powerful than the human beings with disturbing consequences for the future of the humanity. The AI mainly offers domination tools to people who have economical means, politic power to acquire them cognitive skills to use them. AI (eg. profiling, facial recognition, geolocalization...), could be used to enslave the poor and vulnerable populations, thus, main drawback of AI could be to enhance totalitarism in the world.

Transhumanisme

The transhumanisme is an elitist movement which aims at improving the human being by means of the use of the science and the technology to improve its physical characteristics (strength, longevity, disappearance of the suffering, the disease, the ageing and the death) but also mental by amplifying its capacities of perception, of memory and acquisition of knowledge and by deleting the mental disorders. This movement joins in the lineage of the eugenics and constitutes certainly a very dangerous idea for the humanity. This idea confirms the apology of the individual thought and the glorification of the "Me" denounced by Krishnamurti who will bring more suffering to the humanity.

Autonomous weapon

The development of drones and sophisticated autonomous armament systems makes us be afraid that this evolution is inevitable. There are many more weapons than hospitals in our world. An open letter with petition against the autonomous killing robots "Autonomous weapons: the year open letter from HAVE and robotics researchers" Future of Life Institute (2015) was recently displayed online by the researchers in artificial and robotics intelligence to warn the advent of such systems which could fall in the wrong hands and create the chaos at the world level. It seems to me reasonable to sign her and to make her known. The freedom, our lifestyle and even the future of the humanity is at stake. It is our duty of responsibility for warning the disasters bound to the threats which we perceive as seriousness Dupuy (2002) and to warn by throwing this alert. This imperious necessity will maybe make excuse the unusual length of this work.

Conclusion

We described the advantages and drawbacks of memetic and autonomous AI, the hierarchy of information, a classification of the AI systems and the different interactions and roles of human beings that provide and use them. We used as example the AI medical systems that are used in medicine and we compared the consequences of memetic AI and autonomous AI on the patient-caregivers relationship. The state of the art concerning the nature of the thought and the consciousness allowed to confront the arguments in favor and set against the actual feasibility of autonomous system able to act alone and in a second time to be endowed with the capacity of artificial thought towards an artificial consciousness necessarly different from the human consciousness but more ubiquitous and powerful because connected to Internet. The Krishnamurti's first way of thinking that relies on the experience knowledge memory accumulated from humanity history will be implemented soon with AI and Internet, leading to the emergence of an artificial robot personality. The second way of thinking is the individual insight "active-attention" that each of us can live in some occasion, could certainly not be implemented in robots, because it cames from the fact that we belongs to the humanity and share the same philogenetics: we are all cousins. Robots have no history. We presented the medical ethics toolbox that was

extended to study the digital ethics when autonomous systems are implied in people care and the liability consequences in these situations. We also presented the social impact of the AI developement at the population level and at the individual level. Today, most people are connected to Internet and are labelised with their e-reputation which becomes so important that each person who want to be in relation (professionaly, financialy, for business, personaly) with you will have a look on your records on the social networks to get informations concerning your e-reputation. The impact on your life is so important that Big data and autonomous AI, deep learning are building the new digital world where everybody should try to exist. In one hand, AI provides new means to enhance people's life and in the other hand, the powerful AI systems provides means to control and enslave the populations and to destroy democracy. A reflexion on digital epistemology and ethics must be one of the most important challenge of the AI researchers today.

REFERENCES

References

Aamodt A. and Plaza E., 1994. *Case-Based Reasoning, Foundational Issues, Methodological Variations, and System Approaches. AI Communications*, 7(1), 39–59.

Abiteboul S., 2012. *Sciences des données : de la logique du premier ordre à la Toile.* http://books.openedition.org/cdf/529.

Boulanger D. and Colloc J., 1992. *Detecting Heterogeneity in a Multidatabase Environment through an O.O Model.* In *IFIP, DS5, International Conference on Semantics of Interoperable Database Systems, Victoria, Australia.*

Buntine W., 1988. *Generalized Subsumption and Its Applications to Induction and Redundancy. Artificial Intelligence*, 36, 149–176.

Cardon A., 2016. *L'usage des systèmes informatisés pour le contrôle de tout et de tous. Les cahiers du numérique*, 12, no. 1-2, 171–185.

Cardon A., 2018. *Au-delà de l'intelligence artificielle de la conscience humaine à la conscience artificielle.* ISTE.

Carroll L., 1896. *Symbolic Logic, Part I Elementary,.* London Macmillan and Co.

Colloc J., 1985. *Informatique Médicale (SIAMED) Application : Logiciel original d'antibiothérapie médicale.* Med. Phd, Fac. Medicine, University Lyon 1.

Colloc J., 2016. *Perspectives and Ethics of the Autonomous Artificial Thinking Systems. ACTA SYSTEMICA*, 16, no. 1, 31–36. ISSN 1813-4769. Meritorious Scholarly Contribution Award.

Colloc J., 2017. *A Fuzzy Vectorial Space That Avoids to Defuzzify the Membership Functions.* In P.J. Gonçalves (Ed.), *ESM'2017 The 2017 European Simulation and Modelling Conference.* EUROSIS, 13–24.

Colloc J.; Cuvelier M.H.; and Summons P., 2007. *A Clinical Decision Metaknowledge Model.* In *ISC'2007, EUROSIS, Delft, Nederland.* 23–29.

Colloc J. and Lery N., 1989. *Secret Professionnel et Exploitation de Bases de Données Médicales Informatisées.* In *Association des Epidémiologistes de Langue Française. A.D.E.L.F. Bordeaux II.*

Colloc J. and Léry L., 2008. *Un métamodèle d'aide à la decision en éthique médicale. Santé Décision Management*, 11(1-2), 255–274.

Colloc J. and Summons P., 2017. *An approach of the Process of Addiction: A model of the experience.* In D.M. Dubois and G.E. Lasker (Eds.), *Proceedings of the Symposium on Reversible Time, Retardation and Anticipation in Quantum Physics, Biology and Cybernetics 29th Int Conference on Systems Research, Informatics and Cybernetics, Baden-Baden.* IIAS, Vol 1, 103–107.

Dupuy J.P., 2002. *Pour un catastrophisme éclairé Quand l'impossible est certain.* 2-02-066046-6. Editions du Seuil.

Fuchs C., 2003. *Les langues entre Universalisme et relativisme.* In *Le cerveau et la pensée la révolution des sciences cognitives*, Sciences Humaines éditions. 143–151.

Fuchs C., 2004. *Pour Introduire à la linguistique cognitive.* In *La linguistique cognitive*, Ophrys MSH. 1–16.

Future of Life Institute, 2015. *Autonomous weapons: an open letter from AI & robotics researchers.* URL `http://futureoflife.org/open-letter-autonomous-weapons/`.

Galinon-Mélénec B., 2011. *L'homme trace Perspectives anthropologiques des traces contemporaines.* 410. CNRS EDITIONS.

Hofstadter D. and Dennett D., 1981. *The Mind's I.* Basic Books Inc.

Jacquet-Andrieu A. and Colloc J., 2014. *From Self-Awareness to the Consciousness of the "Speaking Subject". International Journal of Computing Anticipatory Systems*, 28, 201–217.

Kolodner J.L., 1983. *Maintaining Organization in a Dynamic Long-Term Memory. Cognitive Science*, 7, no. 4, 243–280.

Kolodner J.L., 1993. *Case-Based Reasoning*. California: Morgan Kaufmann.

Krishnamurti J., 2005. *La nature de la pensée*.

Le Moigne J.L., 1990. *La modélisation des systèmes complexes*. DUNOD-BORDAS.

Léry L. and Colloc J., 2008. *Prise de décision dans l'éthique au quotidien ? Comment décider le soin ? Santé Décision Management*, 11(1-2), 243–254.

Mille A., 2013. *Des traces à l'ère du Web. Intellectica*, 59, 7–28.

Ny J.F.L., 2005. *Comment l'esprit produit du sens*. Odile Jacob.

Ortony A.; Clore G.; and Collins A., 1988. *The cognitive structure of emotions*. Cambridge University press.

Pitrat J., 1990. *Métaconnaissance Futur de l'intelligence artificielle*. HERMES.

Pitrat J., 1993. *L'intelligence artificielle : au-delà de l'intelligence humaine*. In D. Béchillon (Ed.), *Le cerveau : la machine-pensée*, L'Harmattan Paris, 2-7384-1818-X.

Quillian M.R., 1968. *Semantic Memory*. In M. Minsky (Ed.), *Semantic information processing*, Cambridge, MA: MIT Press. 216–270.

Roseman I.J., 1996. *Appraisal determinants of emotions: Constructing a more accurate and comprehensive theory. Cognition & Emotion*, 10, no. 3, 241–278.

Sapir E., 1930. *The Southern Paiute language. Proceedings of the American Academy of Arts and Sciences*, 65, no. 1, 1–296.

Shen Y.; Colloc J.; Jacquet-Andrieu A.; Guo Z.; and Liu Y., 2018. *Constructing Ontology-Based Cancer Treatment Decision Support System with Case-Based Reasoning. CoRR*, abs/1812.01891. URL `http://arxiv.org/abs/1812.01891`.

Shen Y.; Colloc J.; Jacquet-Andrieu A.; and Lei K., 2015. *Emerging Medical Informatics with Case-Based Reasoning for Aiding Clinical Decision in Multi-Agent System. Journal of Biomedical Informatics*, 56, 307–317.

Vitalis A., 1981. *Informatique, pouvoir et libertés*. Economica.

Zighed D.A. A.J. and G. D., 1992. *SIPINA : Méthode et Logiciel*. Ed. Lacassagne.

BIOGRAPHY

Joël Colloc earned his M.D. at the medical faculty of Lyon and a specialty degree of forensic medicine with a degree of clinical toxicology. He received a MSc. degree of IT from the Business School of Lyon (IAE) and a MSc. degree of computer sciences from the engineering school INSA of Lyon. He served as forensic physician at the Edouard Herriot Hospital in the neurological emergency department to cure drug addicted people, medical ethics and developed drug and addiction database. He went on to earn his Ph.D. in computer sciences at the INSA of Lyon. As Hospital assistant at the laboratory of medical computer science and he taught IT at the medical faculty. He was elected as associate professor in computer sciences at IAE of Lyon and he earned his accreditation to supervise researches in sciences at the Lyon 1. He is a Le Havre Normandy University professor in computer sciences since 2003. His main research topics concern e-health and particularly:, fuzzy vectorial spaces (FVS), multi-agent clinical decision support systems (MADSS) and knowledge bases, Case Based Reasoning, ontologies, nervous system modeling and cognitive sciences and AI applications in medicine and human sciences. His human sciences researches try to conciliate the ethics of using Big Data in epidemiological studies, autonomous systems and robots and keeping ethics use of AI in order to improve clinical decision in medicine while preserving the patient-caregiver relationship, the privacy and the freewill choice of the patients.

MULTI-AGENT SOLUTION FOR A DISTRIBUTED INTELLIGENT PHOTO SURVEYING

Anton Ivaschenko, Arkady Krivosheev
Samara State Technical University
244 Molodogvardeyskaya
Samara, Russia
E-mail: anton.ivashenko@gmail.com

Pavel Sitnikov
ITMO University
49 Kronverksky Pr.
Saint Peterburg, Russia
E-mail: sitnikov@o-code.ru

KEYWORDS

Multi-agent Technology, Data Analysis, Image Recognition, Electricity Meter Data Processing.

ABSTRACT

In this paper it is proposed to solve a problem of stream processing of photo images produced by a number of distributed cameras using a multi-agent software solution. Specifically designed intelligent system implementing a neural network is introduced for image recognition. The proposed solution was required to process and analyze the results of electrical meters that are manually monitored by a group of patrol personnel inspectors using hand held devices. This example illustrates a necessity to provide autonomous pre-processing of photo images by an intelligent mobile application with a multi-agent architecture. This paper presents a multi-agent model for this solution and results of its implementation in practice.

INTRODUCTION

Artificial neural networks attract strong interest of scientists and software engineers nowadays. One of the areas of their efficient implementation in practice is image processing, including face detection and verification, medical diagnostics, earth remote sensing, etc. Despite high popularity and propagation of this technology in science, its practical application remains still challenging.

One of perspective areas of neural networks application is an intelligent analysis of photo surveying results. The problem is actual for e.g. electrical power network inspectors that need a mobile tool to capture and process photo and video images that describe the current state of network devices, lines and systems. Although digital devices are widely used nowadays, the number of analog indicators and meters remains high, which forms a problem of their recognition and processing.

This problem can be solved using a network based solution with a front end mobile application for fixation and pre-processing and a database on a server side for storage and analysis. An intelligent component for image recognition can be either server based or deployed on a mobile device: the first option is better for complex images, and the second one provides higher performance.

Both versions require a multi-agent architecture, which was developed for this case and described below.

STATE OF THE ART

Technologies of intelligent pattern recognition using neural networks and positive examples of their practical applications are presented in (Goodfellow, 2016, Egmont-Petersen, 2002). Neural networks provide adequate and stable identification of real objects and items with a complex shape. Textual data (words and numbers) is identified with a sufficiently high accuracy even in case of fuzzy or washed-out picture.

The quality of data processing and analysis can be improved under the context of monitoring procedure, results of preceding identification and multiple photo fixations. Several data sets can be combined to analyze multiple layers of a system at once. This approach is widely used for medical data processing but can be easily disseminated for a cyber-physical system (Holzinger, 2014, 2016). The system can interlink all related data sets (e.g., images, text, measured values, scans) and offer visual analytics to support experts.

Distributed photo surveying can be considered as a problem of the Internet of Things. Combination of the Internet of Things as a major data source and Big Data analysis technologies is a powerful tool for information processing and analysis (Bessis, 2014, Baesens, 2014) and can be successfully used at modern enterprises with distributed structure, i.e. electrical networks and oil pipelines.

Some experience in the area of Big Data processing and analysis for monitoring and maintenance is presented in (Ivaschenko, 2014, 2017). Modern software architectures can help improving the performance of data processing in real time using an approach of parallel computing, multi-agent modelling and distributed decision-making. This fact makes it promise to utilize this technological stack to improve the efficiency of management.

One of the solutions can be close to subject-oriented approach for business processes management (S-BPM), which conceives a process as a collaboration of multiple subjects organized via structured communication (Fleischmann, 2013, 2015). There can be proposed a model for the interaction of actors in integrated information space,

which can be implemented using the multi-agent software (Wooldridge, 2002, Gorodetskii, 2012).

This approach offers a way of designing adaptive systems with decentralization over distributed and autonomous entities organized in hierarchical structures formed by intermediate stable forms. Its implementation in practice requires development of new methods and tools for supporting fundamental mechanisms of self-organization and evolution similar to living organisms (colonies of ants, swarms of bees, etc).

SOLUTION ARCHITECTURE

The Universal Recognizer module was developed for real-time recognition of analog (see Fig. 1) and LCD (see Fig. 2) meters on Android phones. AnyLine OCR application of the Austrian company ANYLINE, created in 2013, is used as an ideological analogue for development. The main convenience of recognizing meterreadings by AnyLine OCR program is that the user just has to turn on the recognition in real time and put a frame on the screen on the meter, the program itself will recognize the type of meter (analog or LCD) and the reading value, after which it will automatically take a photo of the meterand save it together with the value of the readings in the database.

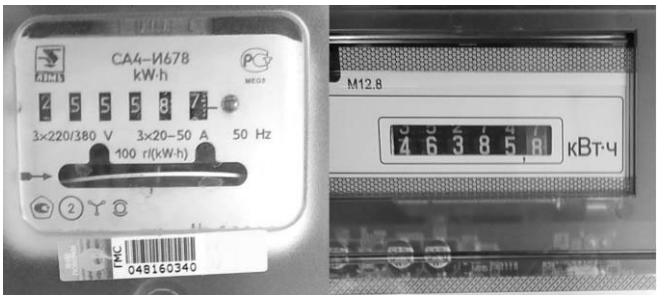


Figure 1: Examples of analog meters

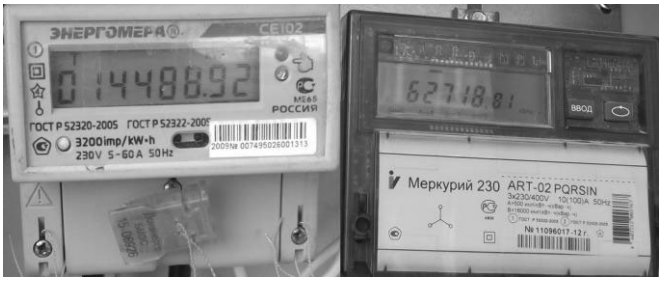


Figure 2: Examples of LCD meters

Figure 3 shows the architecture of the Universal Recognizer module. The logic of the module consists in the following steps:

1. From the camera, each image comes into Controller, and then this image is displayed on the screen.

2. Inside the update controller method performs the following actions:

2.1. Tracker launched is initialized in a separate thread. As initialization parameters, the current image and the frame (the main rectangle visible on the screen by the user) that you want to track are used. If initialization has already occurred, then a new image is transmitted to Tracker, and back, if the result is ready, the updated frame position is returned. If the result is not ready, the frame from the previous image is used. If the position of the updated frame is too different from the original one visible on the screen, then the Controller calls methods that erase all the data currently collected and starts the recognition process again, with the Tracker initialized.

2.2. UniversalRecognizer runs in a separate thread. The first time it is asked for ready to give out the recognition result, if ready, then result is transferred to the Controller by the appropriate method. The second time, it is asked for ready to accept the image for processing, if it is ready, then the image and the updated frame are sent to it.

2.3. If the result was obtained in clause 2.2, then it is transferred to the Agregator. Inside the Agregator, the results are compared with each other and if for a short period of time the results are the same, it will follow that the meterdata was recognized and the corresponding message will be displayed through the Agregator -> Controller -> ViewHelper -> UI chain.

2.4. After receiving the result from section 2.2 (the empty result is also the result), rectangles around the recognized readings are extracted from it and transferred to the ViewHelper. Inside the ViewHelper there is a list of rectangles that are in one of the Activity layers, this allows them to automatically update their position each time the Activity is updated, i.e. on each frame from the camera. This enables the user to observe the recognition process.

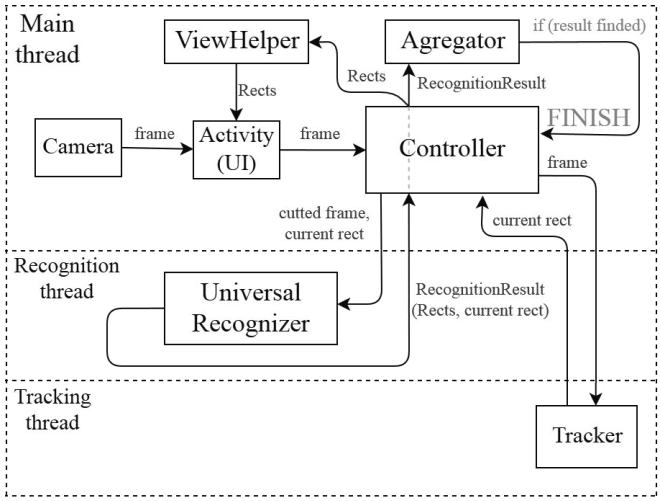


Figure 3: Universal Recognizer architecture

IMPLEMENTATION IN ELECTRICITY METER DATA PROCESSING

Formal statement of the task according to the ToR: to develop a web application for photographing the readings of electricity meters, their transmission to the data processing center, recognition and operational analysis by the staff of one of the Russian regional energy distribution companies.

System Tasks:

- automation of settlements for energy resources with suppliers and consumers;
- automation of the process of recording the readings of electricity meters by employees of the company;
- identification of fraudulent and other illegal actions of consumers;
- collection of evidence of identified offenses for use in judicial and other proceedings.

The system should consist of the following software modules:

- data loading module;
- data recognition module;
- integration module;
- express data analysis module;
- module for generating stickers with QR codes;
- module for constructing routes for bypassing subscribers by employees providing readings from metering devices.

The data recognition module provides the following functions:

1. Recognize the serial number of the meterin accordance with the approved "mask", if available using the factory bar code or QR code;

2. Generate a request (or determine automatically using the available database) to select the type of meter by the number of tariffs (up to 4 values) and control the information received, according to the selected type;

3. Provide self-learning to minimize recognition errors based on experience;

4. In automatic mode, ensure the quality of recognition of meter readings with a liquid crystal meter 100%, meter with a mechanical meter at least 95% of the characters, provided that there are no "transitional values" of the readings and the requirements for photographs of the meter are observed;

Recognition of the QR code (1) was carried out by a program already written for one of the previous projects. Barcode recognition was self-made. Due to the poor quality of the photo (the photo is taken on budget phones, the meters that have stood on the street for many years are photographed, the appearance of such metering devices are far from ideal, etc.) barcode recognition was divided into 3 stages:

1. Finding a barcode in the photo. We searched for areas with a large number of transitions from white to black and vice versa (the horizontal and vertical gradients were considered for vertically pasted barcodes);

2. The found areas were cut out, an attempt was made to restore the barcode;
3. The recovered barcode is recognized by standard means.

Initially, there was an attempt to implement the choice of the type of meter and the number of tariffs (2) by finding and comparing key points with AKAZE algorithms, but due to bad photos, the key points in two photos with meters of the same type differed very much strong. As a result, it was decided to recognize the type of meter, and therefore its number of tariffs, by the arrangement of words on the front side of the meter.

Search for words was carried out by Tesseract. By the relative position of the words, there was a metermask, by the absolute position of the words in the photo there was an area with readings according to the data obtained from the mask. This functionality seems to have even been delivered to the customer, but in the end no one uses it again because of the quality of the photo.

As an alternative, a video stream from the phone's camera was implemented in the browser, a frame was displayed on top of the video stream, the user needed to point the phone at the meterso that the readings fell into the frame, then click on the recognition button.

The self-learning system(3) works as follows: after recognition, if there are incorrect results of digit recognition, the crawler manually corrects the recognized value. The image of the incorrectly recognized digit and the correct value are transmitted to the server. The specially trained userchecks the list of pairs [image - correct value] and adds it to the training set. When a sufficient number of samples are accumulated in the training set, the neural network responsible for recognition is trained. Retraining in this context is to add new images to the entire training set and complete retraining of the neural network on this set.

Recognition (4) works according to the following algorithm. The inspector takes a photo; it is processed by a series of filters that increase the quality and sharpness. Then the colour image is converted to b / w in about 50 times in different ways, each image is subjected to the following actions:

1. Contours are searched for in the image, contours are outlined by rectangles;

2. Among the obtained rectangles, sequences are found that lie on one straight line and having the same size,

among all sequences, the best one is taken according to an empirically selected formula with arguments "rectangle size" and "number of rectangles";

3. According to the rectangles of the b / w image, sections with potential readings are cut out;

4. The cut out images are processed and recognized, the recognition results are saved.

Among the saved results, results are selected for which the recognition percentage is above a certain threshold. Among the selected results is the maximum height of the readings, results on readings with a height well below the maximum are excluded from the sample. The remaining results are combined by location. By the most common recognized value in the combinations, the result is formed.

Recognition of the readings of the LCD and analog meters is different, this is due to the difference in the readings themselves, in the drum in most cases the readings are white on a dark background, the numbers have no gaps, and in the LCD readings are dark segmented values against a light background.

For more details, consider character recognition. For characters from LCD meters, a parallelepiped is located that limits the contour of the potential indication, a mask with cells is applied to the parallelepiped, a parallelepiped with a mask is superimposed on the image, the brightness is calculated in the cells, and the number can be determined from the brightness distribution. Symbols from analogue meters are recognized by the neural network (multilayer perceptron with 2 hidden layers) based on the DeepLearning4j framework. Initially, the neural network was trained on a set of handwritten numeral samples MNIST, the quality of recognition was terrible.

PROBATION RESULTS

The first attempt to improve was to add to the set of MNIST real data obtained during the work of the application from the customer. The data set consisted of ~ 2000 images of digits, mainly zeros and eights. By transformation (shift of the figure relative to the image), the set was increased up to 26,000 images, in the process it turned out that a large number of uniform transformations negatively affects the quality, as a result, the set amounted to 10,000 images. After training, the neural network began to work even worse.

The analysis showed that the neural network is retraining, see Fig. 4, 5. After manually selecting hyper parameters for regularization, the situation with the neural network improved, see Fig. 6, 7, but the quality of recognition of meter readings remained at the same level.

The main problem was associated with the presence in the training set of handwritten numbers that are completely different from the Arabic numbers found in the meter readings. It was necessary to get hold of a set of images of Arabic numerals. To solve the problem, about 1000 fonts were downloaded, based on them, images of numbers were generated, 1000 images per digit were obtained, in the amount of 10,000 source images. By rotation and linear shift, a set of 196,000 images was obtained, divided into a training set (80%) and a test set (20%). It is worth noting that in the original set, the units had almost 30% duplicates. After removing duplicates, an additional rotation angle was used to convert the unit images when converting unit images.

For the obtained set, a lattice search of hyper parameters was applied, therefore, in view of the large number of generated neural networks; data for constructing graphs were not collected. The search resulted in a neural network with quality:

- measure F1 on the training set = 0.9943;
- measure F1 on the test set = 0.9903;
- standard error on the training set = 0.0282;
- standard error on the test set = 0.0368.

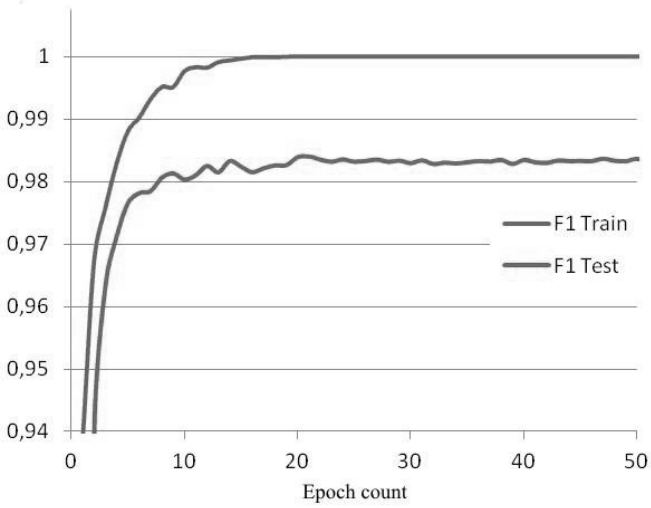


Figure 4: Neural network training results (F1)

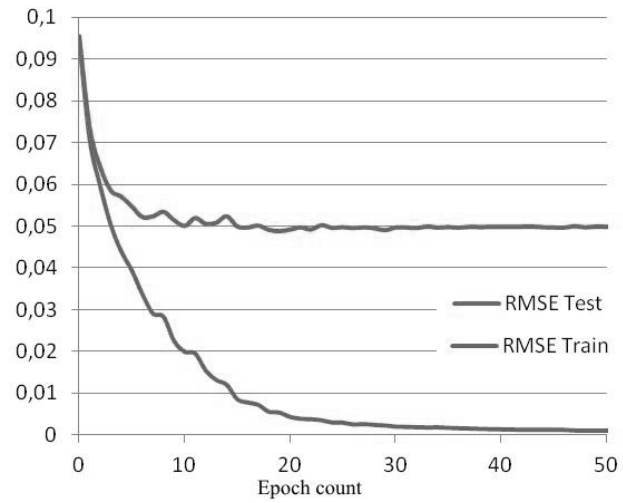


Figure 5: Neural network training results (RMSE)

The quality of recognition of indications when using the new neural network immediately increased. For a numerical assessment of recognition accuracy, tests were written. A set of photographs of analog and LCD meters was collected, which included photographs of very different quality, each photograph had a dimensioned area with indications (the coordinates of the describing rectangle were found); a photograph was sent to the UniversalRecognizer and 10 different options for frame.

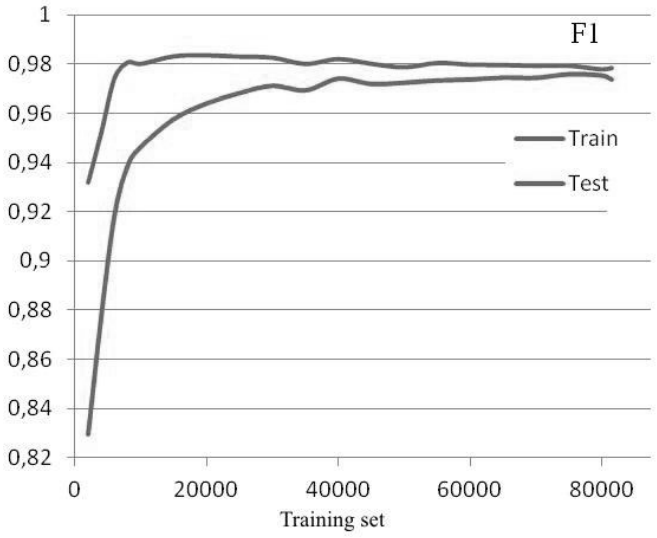


Figure 6: Identification improved (F1)

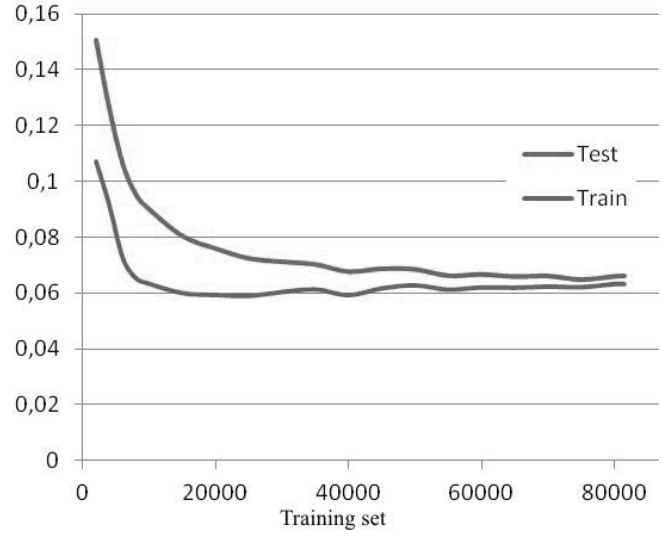


Figure 7: Identification improved (RMSE)

Accuracy was calculated according to the worst cases of isolation, when a minimum of indications was recognized, the best and on average for all emissions. As a result, 138 images of LCDs (total 777 digits) of meters and 95 images of analog meters (total 534 digits) were collected, recognition results:

- Analog meters:
  - recognized minimum: 337/534 = 63.11%
  - recognized on average: 432.1 / 534 = 80.92%
  - maximum recognized: 469/534 = 87.83%
- LCD meters:
  - recognized minimum: 427/777 = 54.95%
  - recognized on average: 640.9 / 777 = 82.48%
  - recognized maximum: 758/777 = 97.55%

If the sort of photo set by quality and choose the best half, the recognition results will be as follows:

- Analog meters (42 pcs.):
  - recognized minimum: 196/212 = 92.45%
  - recognized on average: 203.2 / 212 = 95.85%
  - recognized maximum: 209/212 = 98.58%
- LCD meters (70 pcs.):
  - recognized minimum: 368/378 = 97.35%
  - recognized on average: 373.8 / 378 = 98.89%
  - recognized maximum: 378/378 = 100.00%

The algorithm execution time in one photo and one frame on a budget phone is on average from 1 to 5 seconds.

CONCLUSION

Results of the proposed solution implementation in energy company prove the efficiency of multi-agent architecture being used for a distributed processing of image data generated in the process of photo surveying. Next steps include changing the architecture of the solution by using a convolutional neural network and application of this solution to solve the problems of other types of images processing required by energy networks monitoring.

REFERENCES

Baesens, B. 2014. “Analytics in a Big Data world: The essential guide to data science and its applications”. Wiley, 232 p.

Bessis, N.; and C. Dobre. 2014. “Big Data and Internet of Things: A roadmap for smart environments”, *Studies in computational intelligence*, Springer, 450

Egmont-Petersen, M.; D. de Ridder, H. Handels. 2002. “Image processing with neural networks – a review”. *Pattern Recognition*. 35 (10). 2279-2301

Goodfellow, I.; Y. Bengio, A. Courville, and Y. Bengio. 2016. “Deep learning, vol. 1” MIT press Cambridge, Vol. 925

Gorodetskii, V.I.. 2012. “Self-organization and multiagent systems: I. Models of multiagent self-organization”, *Journal of Computer and Systems Sciences International*, vol. 51, issue 2. 256-281

Fleischmann, A.; U. Kannengiesser, W. Schmidt, and C. Stary. 2013. “Subject-oriented modeling and execution of multi-agent business processes”. *In Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, 2013 IEEE/WIC/ACM International Joint Conferences, 138-145

Fleischmann, A.; W. Schmidt, and C. Stary. 2015. “S-BPM in the wild”. Springer, 282

Holzinger, A. 2014. “Extravaganza tutorial on hot ideas for interactive knowledge discovery and data mining in biomedical informatics”. *Lecture Notes in Computer Science*, 8609, 502-515

Holzinger, A. 2016. “Interactive machine learning for health informatics: when do we need the human-in-the-loop?”. *Brain Informatics*, Volume 3, Issue 2, 119-131

Ivaschenko, A.; M. Milutkin, and P. Sitnikov. 2017. “Accented visualization in maintenance AR guides”. *Proceedings of SCIFI-IT 2017 Conference*, Belgium, EUROSIS-ETI. 42-45

Ivaschenko A.; A. Khorina, V. Isayko, D. Krupin, and P. Sitnikov. 2018. “Online creativity modeling and analysis based on Big Data of social networks”. *Proceedings of Computing Conference 2018*, London, UK. 199-202

Wooldridge, M. 2002. “An introduction to multi-agent systems”, John Wiley and Sons, Chichester. 340 p.

FAULT DETECTION OF ELEVATOR SYSTEM USING PROFILE EXTRACTION AND DEEP AUTOENCODER FEATURE EXTRACTION

Krishna Mohan Mishra[1], John-Eric Saxen[2], Jerker Bjorkqvist[2], and Kalevi J. Huhtala[1]
[1]Unit of Automation Technology and Mechanical Engineering, Tampere University, Tampere, Finland
[2]Department of Information Sciences, Abo Akademi University, Turku, Finland
email: {krishna.mishra, kalevi.huhtala}@tuni.fi
email: {John-Eric.Saxen, jerker.bjorkqvist}@abo.fi

KEYWORDS
Elevator system, deep autoencoder, fault detection, feature extraction, random forest, profile extraction

ABSTRACT

In this paper, we propose a new algorithm for data extraction from time series data, and furthermore automatic calculation of highly informative deep features to be used in fault detection. In data extraction elevator start and stop events are extracted from sensor data, and a generic deep autoencoder model is also developed for automated feature extraction from the extracted profiles. After this, extracted deep features are classified with random forest algorithm for fault detection. Sensor data are labelled as healthy and faulty based on the maintenance actions recorded. The rest of the healthy data are used for validation of the model to prove its efficacy in terms of avoiding false positives. We have achieved nearly 100% accuracy in fault detection along with avoiding false positives based on new extracted deep features, which outperforms results using existing features. Existing features are also classified with random forest to compare results. Our developed algorithm provides better results due to the new deep features extracted from the dataset when compared to existing features. This research will help various predictive maintenance systems to detect false alarms, which will in turn reduce unnecessary visits of service technicians to installation sites.

INTRODUCTION

In recent years, elevator systems have been used more and more extensively in apartments, commercial facilities and office buildings. Nowadays 54% of the world's population lives in urban areas (Desa 2014). Therefore, elevator systems need proper maintenance and safety. Fault diagnosis methods based on deep neural networks (Jia et al. 2016) and convolutional neural networks (Xia et al. 2018) feature extraction methodology are presented as state of the art for rotatory machines similar to elevator systems. Support vector machines (Martínez-Rego et al. 2011) and extreme learning machines (Yang and Zhang 2016) are also used as fault detection methods for rotatory machines. However, we have developed an intelligent deep autoencoder random forest based feature extraction methodology for fault detection in elevator systems to improve the performance of traditional fault diagnosis methods.

Acceleration profile extraction methods have applied in electric vehicles (Bingham et al. 2012) and horizontal planes (Soyka et al. 2011). Kalman filter (Wang et al. 2015) is one of the methods being used for acceleration profile extraction. However, we have developed an offline profile extraction algorithm based on low-pass filtering and peak detection to extract elevator start and stop events from sensor data.

Autoencoders were first introduced by LeCun (Fogelman-Soulie et al. 1987), and have been studied for decades. Traditionally, feature learning and dimensionality reduction are the two main features of autoencoders. Recently, autoencoders have been considered one of the most compelling subspace analysis techniques because of the existing theoretical relations between autoencoders and latent variable models. Autoencoders have been used for feature extraction from the data in systems like induction motor (Sun et al. 2016) and wind turbines (Jiang et al. 2018) for fault detection, different from elevator systems as in our research.

In our previous research, raw sensor data, mainly acceleration signals, were used to calculate elevator key performance and ride quality features, which we call here existing features. Random forest was used for fault detection based on these existing features. Existing domain specific features are calculated from raw sensor data, but that requires expert knowledge of the domain and results in a loss of information to some extent. To avoid these implications, we have developed an algorithm for profile extraction from the raw sensor data rides and a generic algorithm with deep autoencoder random forest approach for automated feature extraction from raw sensor data profiles for fault detection in elevator systems. The rest of this paper is organized as follows. The next section presents the methodology of the paper including profile extraction, deep autoencoder and random forest algorithms. This is followed

by section that includes the details of experiments performed, results and discussion. Finally, the last section concludes the paper and presents the future work.

METHODOLOGY

In this study, we have utilised 12 different existing features derived from raw sensor data describing the motion and vibration of an elevator for fault detection and diagnostics of multiple faults. We have developed an automated feature extraction technique for raw sensor data in this research as an extension to the work of our previous research (Mishra et al. 2019) to compare the results using new extracted deep features. We only extract start and stop profiles from the rides because of the different lengths of rides for each floor combination due to the constant speed phase, which is longer when there is longer travel.

Profile extraction algorithm

The algorithm works in two stages. In the first stage, the signal is pre-processed and normalized, followed by low-pass filtering in order to reduce noise spikes. The low-pass filtered signal is used for peak detection, which for each elevator travel detects a local minimum and maximum, corresponding to acceleration and deceleration (start and stop) events.

In the second stage, alignment and collection of equal length profiles is performed based on windowing of the acceleration signal near the peak events. In this stage, the raw acceleration signal is used instead of the filtered signal. A number of time domain alignment methods have been proposed in the literature. Dynamic time warping (DTW) has been commonly applied, e.g. in speech recognition (Di Martino 1985), whereas various alignment techniques for sensor data have been presented in (Rhudy 2014). Here, alignment is performed against a reference profile, which is initialized to the known approximate length of the acceleration and deceleration windows. The reference profile is aligned against the raw data in the window of the detected peaks. The criterion for optimal alignment was defined as the alignment that minimizes the sum of the Euclidean or L_2 norm. The output from this operation is an $n \times m$ matrix of aligned profiles describing n acceleration and deceleration events of length m.

In order to improve the alignment accuracy, the reference profile is updated iteratively following each run. Each sequence in the profile matrix is of the same sample size and closely synchronized in time and can hence be considered a repetition of the same signal. Using signal averaging, the new reference profile is calculated as the mean of the n extracted profiles. This both maintains the main characteristics of the signal and reduces the noise. Assuming white noise and perfect synchronization, signal averaging improves the signal-to-noise ratio *(SNR)* by a factor of $\sqrt{n}$. The reference profile is updated on-line during the alignment stage or in batch mode by multiple iterations through the same dataset. The off-line profile extraction algorithm is described as follows.

Off-line profile extraction algorithm

Pre-procession

1. Read a vector of raw acceleration data containing k elevator travels. Define the zero mean transformed dataset as X.
2. Perform low-pass filtering on X and obtain denoised dataset Y.

Initialization

3. Define parameters for reference profile. Set window length to m samples and height h to the 99th percentile of the low-pass filtered dataset.
4. Set threshold limit t for triggering peak detection as a fraction of h.
5. Define alignment window size a and set k=1.

Iteration

6. From *Y(k)*, detect peak acceleration points y_{min} and y_{max} satisfying $abs(y_{min,max}) \geq t$
7. Align reference profile R against raw dataset X in the vicinity of detected peaks by minimizing the L_2 norm according to

$$min \sum_{i=-a/2}^{a/2} \sum_{j=1}^{m} [-r_{\mathrm{j}} - x_{\mathrm{min+i+j}}]^2 \quad (1)$$

$$min \sum_{i=-a/2}^{a/2} \sum_{j=1}^{m} [r_{\mathrm{j}} - x_{\mathrm{max+i+j}}]^2 \quad (2)$$

8. Add aligned data points from *X(k)* as rows into an $n \times m$ profile matrix, alternatively separate matrices according to direction of travel (min/max).
9. Set travel window k=k+1 and repeat steps 6-8 until end of dataset.
10. Update reference profile with the signal-averaged profile obtained from the column-wise mean of the new profile matrix. Set k=1 and continue with new batch iterations by repeating steps 6-9.

Deep autoencoder

We are using a five layer deep autoencoder (see Figure 1) including input, output, encoder, decoder and representation layers, which is a different approach than in (Jiang et al. 2018), (Vincent et al. 2008). In our approach, we first analyze the data to find all floor patterns and then feed the segmented raw sensor data windows in up and down directions separately to the algorithm for profile extraction. Extracted profiles are fed to the deep autoencoder model for extracting new deep features. Lastly, we apply random forest as a classifier for

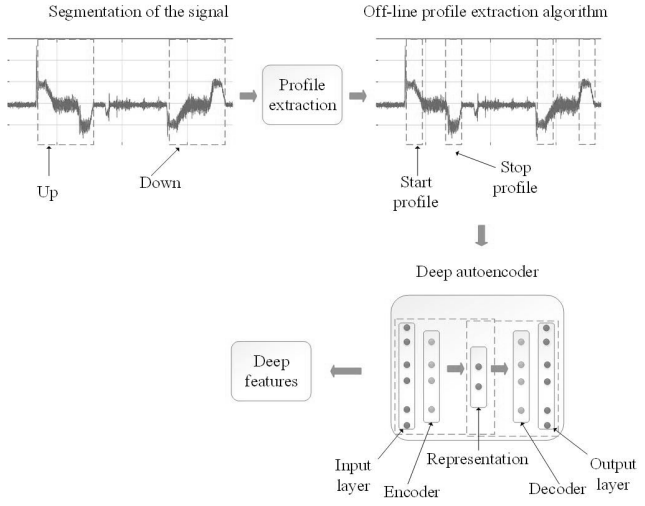


Figure 1: Off-line profile extraction and deep autoencoder feature extraction approach.

fault detection based on new deep features extracted from the profiles.

The encoder transforms the input x into corrupted input data $x^{'}$ using hidden representation H through nonlinear mapping

$$H = f(W_1 x^{'} + b) \tag{3}$$

where *f(.)* is a nonlinear activation function as the sigmoid function, $W_1 \in R^{k*m}$ is the weight matrix and $b \in R^k$ the bias vector to be optimized in encoding with k nodes in the hidden layer (Vincent et al. 2008). Then, with parameters $W_2 \in R^{m*k}$ and $c \in R^m$, the decoder uses nonlinear transformation to map hidden representation H to a reconstructed vector $x^{"}$ at the output layer

$$x^{"} = g(W_2 H + c) \tag{4}$$

where *g(.)* is again nonlinear function (sigmoid function). In this study, the weight matrix is $W_2 = {W_1}^T$, which is tied weights for better learning performance.

Random forest

The final classification accuracy of random forest is calculated by averaging, i.e. arithmetic mean of the probabilities of assigning classes related to all the produced trees *(e)*. Testing data *(d)* that is unknown to all the decision trees is used for evaluation by the voting method. Specifically, let sensor data value ${v_l}^e$ have training sample l^{th} in the arrived leaf node of the decision tree $e \in E$, where $l \in [1, ..., L_e]$ and the number of training samples is L_e in the current arrived leaf node of decision tree e. The final prediction result is given by (Huynh et al. 2016):

$$\mu = \frac{\sum_{e \in E} \sum_{l \in [1,...,L_e]} {v_l}^e}{\sum_{e \in E} L_e} \tag{5}$$

All classification trees providing a final decision by voting method are given by:

$$H(a) = \arg\max_{y_j} \sum_{i \in [1,2,...,Z]} I(h_i(a) = y_j) \tag{6}$$

where *j= 1,2,...,C* and the combination model is *H(a)* , the number of training subsets are Z depending on which decision tree model is $h_i(a)$, $i \in [1, 2, ..., Z]$ while output or labels of the P classes are y_j , *j= 1,2,...,P* and combined strategy is *I(.)* defined as:

$$I(x) = \begin{cases} 1, & h_i(a) = y_j \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

where output of the decision tree is $h_i(a)$ and i^{th} class label of the P classes is y_j, *j= 1,2,...,P*.

RESULTS AND DISCUSSION

In this research, we first selected all floor patterns like floor 2-5, 3-8 and so on from the data, some of which are shown in Table 1.

Table 1: Floor patterns.

Start floor	Stop floor
0	1
2	5
3	8
4	6

The next step includes the selection of faulty rides from all floor patterns based on time periods provided by the maintenance data. An equal amount of healthy rides are also selected. Only the vertical component of acceleration data is selected in this research because it is the most informative aspect, consisting of significant changes in vibration levels as compared to other components. Healthy and faulty rides are fed to the algorithm for profile extraction separately. Start and stop profiles are of equal length, irrespective of floor combination.

Up movement

We have analyzed up and down movements separately because the traction based elevator usually produces slightly different levels of vibration in each direction. First, we have selected faulty rides based on time periods provided by the maintenance data, including all floor patterns, which is fed to the algorithm for profile extraction, as shown in Figure 2.

Then we have selected an equal number of rides for healthy data, similar to Figure 2. The next step is to

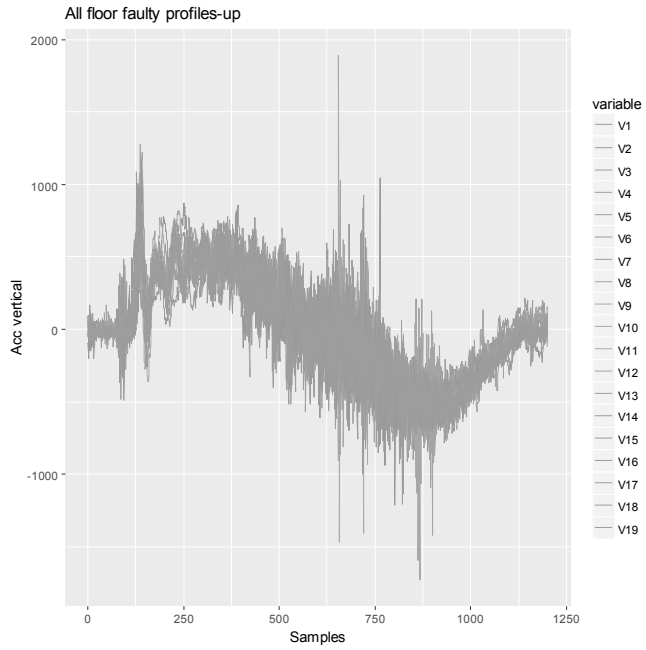


Figure 2: Profiles from faulty rides (Acc represents acceleration signal).

label both the healthy and faulty profiles with class labels 0 and 1 respectively. Healthy and faulty profiles with class labels are fed to the deep autoencoder model and the generated deep features are shown in Figure 3. These are called deep features or latent features in deep autoencoder terminology, which shows hidden representations of the data.

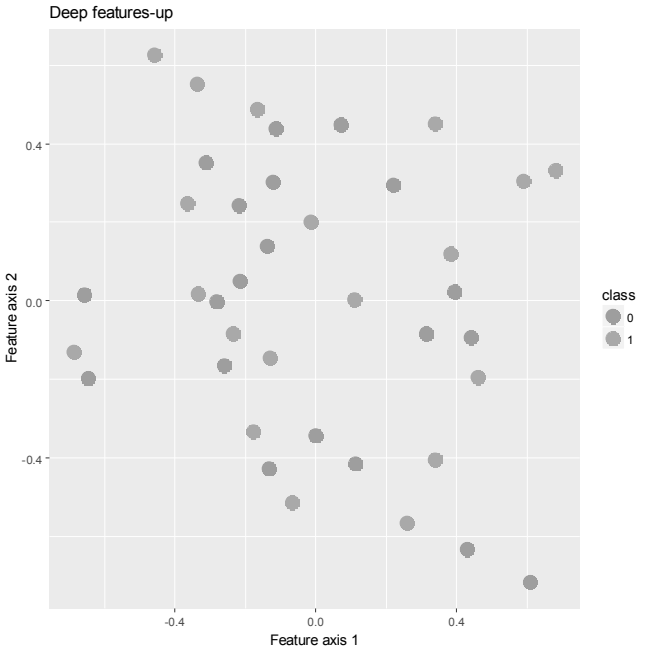


Figure 3: Extracted deep autoencoder features (visualization of the features w.r.t class variable).

Extracted deep features are fed to the random forest algorithm for classification, and the results provide 100% accuracy in fault detection in Table 2. We have compared accuracy in terms of avoiding false positives from both features and found that new deep features generated in this research outperform the existing features. We have used the rest of the healthy rides for extracting profiles to analyze the number of false positives. These healthy profiles are labelled as class 0 and fed to the deep autoencoder to extract new deep features from the profiles.

These new deep features are then classified with the pretrained deep autoencoder random forest model to test the efficacy of the model in terms of false positives. Table 2 presents the results for upward movement of the elevator in terms of accuracy, sensitivity and specificity. We have also included the accuracy of avoiding false positives as an evaluation parameter for this research. The results show that the new deep features provide better accuracy in terms of fault detection and avoiding false positives from the data, which is helpful in detecting false alarms for elevator predictive maintenance strategies. It is extremely helpful in reducing unnecessary visits by maintenance personnel to installation sites.

Table 2: Fault detection analysis (False positives field related to analyzing rest of the healthy profiles after the training and testing phase).

	Deep features	**Existing features**
Accuracy	1	0.55
Sensitivity	1	0.33
Specificity	1	0.80
False positives	1	0.48

Down movement

For downward motion, we have repeated the same analysis procedure as in the case of upward motion. Table 3 presents the results for fault detection with deep autoencoder random forest model in the downward direction. The results are similar to the upward direction but we can see significant change in terms of accuracy of fault detection and when analyzing the number of false positives with new deep features.

CONCLUSIONS AND FUTURE WORK

This research focuses on the health monitoring of elevator systems using a novel fault detection technique. The goal of this research was to develop generic models for profile extraction and automated feature extrac-

Table 3: Fault detection analysis.

	Deep features	Existing features
Accuracy	1	0.78
Sensitivity	1	0.60
Specificity	1	1
False positives	0.98	0.66

tion for fault detection in the health state monitoring of elevator systems. Our approach in this research provided nearly 100% accuracy in fault detection and also in the case of analyzing false positives for all floor combinations with new extracted deep features. The results support the goal of this research of developing generic models which can be used in other machine systems for fault detection. Our models outperform others because of new deep features extracted from the dataset as compared to existing features calculated from the same raw sensor dataset. The automated feature extraction approach does not require any prior domain knowledge. It also provides dimensionality reduction and is robust against overfitting characteristics.
In future work, we will extend our approach on more elevators and real-world big data cases to validate its potential for other applications and improve its efficacy.

REFERENCES

Bingham C.; Walsh C.; and Carroll S., 2012. *Impact of driving characteristics on electric vehicle energy consumption and range. IET Intelligent Transport Systems*, 6, no. 1, 29–35.

Desa, 2014. *World urbanization prospects, the 2011 revision. Population Division, Department of Economic and Social Affairs, United Nations Secretariat.*

Di Martino J., 1985. *Dynamic time warping algorithms for isolated and connected word recognition.* In *New systems and architectures for automatic speech recognition and synthesis*, Springer. 405–418.

Fogelman-Soulie F.; Robert Y.; and Tchuente M., 1987. *Automata networks in computer science: theory and applications.* Manchester University Press and Princeton University Press.

Huynh T.; Gao Y.; Kang J.; Wang L.; Zhang P.; Lian J.; and Shen D., 2016. *Estimating CT image from MRI data using structured random forest and auto-context model. IEEE transactions on medical imaging*, 35, no. 1, 174.

Jia F.; Lei Y.; Lin J.; Zhou X.; and Lu N., 2016. *Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. Mechanical Systems and Signal Processing*, 72, 303–315.

Jiang G.; Xie P.; He H.; and Yan J., 2018. *Wind turbine fault detection using a denoising autoencoder with temporal information. IEEE/ASME Transactions on Mechatronics*, 23, no. 1, 89–100.

Martínez-Rego D.; Fontenla-Romero O.; and Alonso-Betanzos A., 2011. *Power wind mill fault detection via one-class ν-SVM vibration signal analysis.* In *Neural Networks (IJCNN), The 2011 International Joint Conference on.* IEEE, 511–518.

Mishra K.M.; Krogerus T.; and Huhtala K., 2019. *Fault detection of elevator systems using deep autoencoder feature extraction.* In *Research Challenges in Information Science (RCIS), 2019 International Conference on.* IEEE, 43–48.

Rhudy M., 2014. *Time alignment techniques for experimental sensor data. Int J Comput Sci Eng Survey*, 5, no. 2, 1–14.

Soyka F.; Giordano P.R.; Beykirch K.; and Bülthoff H.H., 2011. *Predicting direction detection thresholds for arbitrary translational acceleration profiles in the horizontal plane. Experimental brain research*, 209, no. 1, 95–107.

Sun W.; Shao S.; Zhao R.; Yan R.; Zhang X.; and Chen X., 2016. *A sparse auto-encoder-based deep neural network approach for induction motor faults classification. Measurement*, 89, 171–178.

Vincent P.; Larochelle H.; Bengio Y.; and Manzagol P.A., 2008. *Extracting and composing robust features with denoising autoencoders.* In *Proceedings of the 25th international conference on Machine learning.* ACM, 1096–1103.

Wang X.; Wu S.; and Wang K., 2015. *A run-to-run profile control algorithm for improving the flatness of nano-scale products. IEEE Transactions on Automation Science and Engineering*, 12, no. 1, 192–203.

Xia M.; Li T.; Xu L.; Liu L.; and de Silva C.W., 2018. *Fault Diagnosis for Rotating Machinery Using Multiple Sensors and Convolutional Neural Networks. IEEE/ASME Transactions on Mechatronics*, 23, no. 1, 101–110.

Yang Z.X. and Zhang P.B., 2016. *ELM Meets RAE-ELM: A hybrid intelligent model for multiple fault diagnosis and remaining useful life predication of rotating machinery.* In *Neural Networks (IJCNN), 2016 International Joint Conference on.* IEEE, 2321–2328.

PREDICTING MOTOR POLICY LOSS - A ZAIG MODEL OR A TWO-STAGE NEURAL NETWORK APPROACH?

Luke Aarohi and David Suda
Department of Statistics and Operations Research
University of Malta
Msida MSD2080
Malta
E-mail: luke.aarohi.15@um.edu.mt, david.suda@um.edu.mt

KEYWORDS
Insurance, motor claims, neural networks, zero-adjusted inverse Gaussian model.

ABSTRACT

Artificial neural networks have increasingly being applied to solve problems which traditionally would have fallen under the domain of more classical statistical methodology, and the latter has long been a staple of popular actuarial methodology. We aim to compare a two-stage artificial neural network approach with the zero-adjusted inverse Gaussian model for predicting the claim of a motor insurance policy, which is a popular method with actuaries. The performance of both approaches is analysed by means of *K*-fold cross-validation. The conclusion reached is that our approach provides a comparable, if not superior, overall performance in predicting policy loss which is more robust to extreme observations.

INTRODUCTION

The cardinal concept behind insurance is the protection of the insured members against a potentially hazardous future random event, which is achieved though the creation of a fund by the insurer to which the insured members contribute through their premiums. The positive aspect of the formation of such a communal fund is that the amalgamation of policy holder's risks leads to the reduction of the individual risk. As for any business, an insurance company can only be profitable if the income exceeds expenses. However, the insurance company must also aim to make its premiums competitive. Thus, being able to quantify risk is important for the business. For insurers to be in a position to settle future claims, it is critical that they adopt an effective policy pricing strategy by correctly modelling historical policy data, with the aim of predicting policy loss. Total policy loss of an insurance company in an upcoming period can also be projected, with the aim of setting aside adequate reserves. To calculate the predicted policy loss for an individual policy, we require to forecast the following: (i) the probability of claim occurrence; and (ii) the severity of claims if they occur. This must be based on a number of features of the policy holder and the asset being insured. It must be noted here that the predicted policy loss will be an 'average' (though not always in the mathematical sense) of what actually happens in reality, which is either no claim at all or losses based on some distribution. Indeed, the predicted policy loss we talk about takes the role of a 'break-even premium' for an individual policy, i.e. a premium which, on average, yields a profit/loss of zero. Break-even premiums, though not feasible business-wise, provide a lowerbound to what makes an acceptable gross (charged) premium to an insurance company. They can also serve as a basis to calculate gross premiums. In this paper we shall deal specifically with the motor branch of insurance. As is well-known, this kind of insurance is mandatory, and therefore gives insurance companies a substantial amount of data for modelling claims.

The aim of this paper is that of seeing whether a two-stage artificial neural network (ANN) approach can be an improvement over statistical modelling in claim prediction for an insurance policy. A number of statistical approaches can be found (e.g. Brockman and Wright 1993, Haberman and Renshaw 1996, Pinquet 1997). In these cases, the modelling of claim occurrence and claim severity was conducted by formulating two separate models. Two approaches which model these simultaneously are the Tweedie model (Jorgensen and Paes De Souza 1994, Smyth and Jorgensen 2002), and the zero-adjusted inverse Gaussian (ZAIG) model (Heller et al. 2006). The latter has been highly touted by academics within the field of actuarial science as the most appropriate statistical model for predicting policy loss of insurance policies (see. e.g. Bortoluzzo et al. 2011, Noriszura 2013). ANNs, on the other hand, have been applied to a range of different actuarial scenarios. Similar papers to our approach we find one which proposes the use of ANNs are used to predict the total annual claim amount an insurance company could expect to incur in an upcoming year (Dalkilic et al. 2009), and one where claim frequency and claim severity are predicted separately via two different neural networks (Yunos et al. 2016). To date, we have not found an ANN-based approach which serves the same purpose as the ZAIG model, and our aim is thus to propose one. Other not directly related approaches to modelling claim counts and/or claim sizes can be found, for example where a fully Bayesian approach is conducted which allows for dependency between number of claims and clime size (Gschlößl and Czado 2008), where an ensemble modelling approach is taken for predicting motor insurance claims (Ye et al. 2018, preprint), and where the XGBoost algorithm is implemented for predicting motor insurance claim occurrence and compared to logistic regression (Pesantez-Narvaez et al. 2019).

We adopt a two-stage approach, where we model claim probability (rather than frequency) on the whole dataset and annual claim severity on the claimants' policies, to finally come up with the predicted policy loss for any set of input variables. Since the aim of our two-stage neural network approach is that of creating an analogous alternative to the

ZAIG model by (Heller et al. 2006), we ultimately intend to compare the outcome of this approach on unseen data with that of the ZAIG model. The next two sections of this paper will be dedicated to the introduction of these modelling approaches.

THE ZAIG MODEL

The ZAIG model is constructed as follows. Let Y_i be the size of the claim for the i^{th} policy for $i = 1, \dots N$. We may denote the distribution of Y_i through the following mixed discrete-continuous probability density function:

$$f(y_i) = \begin{cases} 1 - \pi_i \text{ if } y_i = 0 \\ \pi_i \dfrac{1}{\sqrt{2\pi y_i^2 \sigma_i^2}} \exp\left[-\dfrac{1}{2y_i}\left(\dfrac{y_i - \mu_i}{\mu_i \sigma_i}\right)\right] \text{ if } y_i > 0 \end{cases}, \tag{1}$$

where π_i is the probability of experiencing a claim, and the term to the right of π_i in the second line represents the inverse Gaussian density with parameters μ_i and σ_i. The parameters μ_i, σ_i and π_i are indexed by i because they depend on other characteristics (inputs) pertaining to the i^{th} policy. μ_i and π_i are of particular interest to us as they represent mean claim severity and claim probability respectively, while the addition of σ_i yields the inverse Gaussian distribution for claim severity. We incorporate these inputs (consisting of covariates and factor dummy variables) into our model by means of a linear predictor in the same way we would do if we were dealing with regular generalized linear models. The link functions for the linear predictors of the parameters μ_i, σ_i and π_i are

$$\log(\mu_i) = \mathbf{x}'_{\mu_i}\boldsymbol{\beta}_\mu \,,$$
$$\log(\sigma_i) = \mathbf{x}'_{\sigma_i}\boldsymbol{\beta}_\sigma$$
$$\text{logit}(\pi_i) = \log\frac{\pi_i}{1-\pi_i} = \mathbf{x}'_{\pi_i}\boldsymbol{\beta}_\pi \,.$$

In these equations, $\mathbf{x}_{\mu_i}$, $\mathbf{x}_{\sigma_i}$ and $\mathbf{x}_{\pi_i}$ refer to the i^{th} vectors of observed values of the covariates for μ_i, σ_i and π_i respectively. These vectors each consist of a subset of the set of elements within the overall i^{th} vector of inputs $\mathbf{x}_i$. We may opt to use different inputs to model μ_i, σ_i and π_i. Given properties of the mixed discrete-continuous probability density in (1), we have that

$$E[Y_i] = \pi_i \mu_i$$

and

$$Var[Y_i] = \pi_i \mu_i^2 (1 - \pi_i + \mu_i \sigma_i^2).$$

$E[Y_i]$, the expectation, may be used as the prediction $\hat{y}_i$ of the i^{th} policy loss for the ZAIG model.

It is important to note that the ZAIG distribution is not part of the exponential family of distributions and, as a result, the estimation of parameters cannot be carried out by directly applying the maximum likelihood approach used when dealing with generalized linear models. In this case, the maximum penalized likelihood estimation approach is used (Rigby and Stasinopoulos 2005). Moreover, just as with many likelihood-based statistical models, AIC and BIC can be used to find the model of best fit within this framework. We now move on to the next section, where we construct a two-stage ANN approach for estimating the expected claim size.

A TWO-STAGE NEURAL NETWORK APPROACH

Consider an ANN with with depth $k + 1$, hence consisting of k hidden layers. Let $\mathbf{x}$ be a vector of p inputs for any given observation. Then the n_1 activations $\mathbf{a}^{(1)}$ of the first hidden layer are given by

$$\mathbf{a}^{(1)} = f^{(1)}\left(\mathbf{W}^{(1)\prime}\mathbf{x} + \mathbf{b}^{(1)}\right),$$

where $\mathbf{b}^{(1)}$ is the vector of n_1 bias terms, $\mathbf{W}^{(1)}$ is a $p \times n_1$ weight matrix and $f^{(1)}$ is the first activation function. Similarly for $j = 2, \dots, k$ we have

$$\mathbf{a}^{(j)} = f^{(j)}\left(\mathbf{W}^{(j)\prime}\mathbf{a}^{(j-1)} + \mathbf{b}^{(j)}\right),$$

where $\mathbf{b}^{(j)}$ is the vector of n_j bias terms, $\boldsymbol{W}^{(j)}$ is an $n_j \times n_j$ weight matrix and $f^{(j)}$ is the j^{th} activation function. Note that the dimension of $\mathbf{a}^{(j)}$ represents the number of neurons in hidden layer j. Finally the predicted o-dimensional output of the ANN, given by $\hat{y}$ is given by

$$\hat{y} = f^{(k+1)}\left(\mathbf{W}^{(k+1)\prime}\mathbf{a}^{(k)} + \mathbf{b}^{(k+1)}\right),$$

where $\mathbf{b}^{(k+1)}$ is the vector of o bias terms, $\mathbf{W}^{(k+1)}$ is an $n_k \times o$ weight matrix and $f^{(k+1)}$ is the $(k+1)^{\text{th}}$ activation function.

We now discuss the use of activation functions for the hidden and output layers. We focus our efforts on the choice between two activation functions for the hidden layers - the rectified linear unit (ReLU) and the leaky ReLU (Glorot 2001, Maas et al. 2013). At each level of the input, the ReLU function takes the form $f(\mathbf{x}) = \mathbf{x}^\vee$, where $\mathbf{x}^\vee$ has the same positive components as $\mathbf{x}$ with the negative components replaced by 0. On the other hand, the leaky ReLU function takes the form $f(\mathbf{x}) = \mathbf{x}^\vee + \alpha\mathbf{x}^\wedge$ where α is fixed to be a very small value such as 0.01 and $\mathbf{x}^\wedge$ has the same negative components as $\mathbf{x}$ with the positive components replaced by 0. Finally, at the output layer, the choice of the activation function is typically taken to be the identity in the case of regression problems, the logistic sigmoid function in the case of binary classification problems, and the softmax function in multiclass problems (see, e.g. Goodfellow et al. 2015), with $o = 1$ in the first and second case and $o > 1$ in the third case. In our context, we take the logistic sigmoid function to model claim probability, which is a binary classification problem, and we take the identity to model log claim severity on the reduced claimant dataset, thus having a regression problem. Finally, we note that the mini-batch stochastic gradient descent (or mini-batch SGD) is the algorithm behind the parameter estimation process of ANNs, which depends on two pre-specified parameters: the learning rate of the iterative process and the batch sizes into which the dataset is partitioned.

Goodness-of-fit of ANNs can be performed using an overall loss function. Given output $\mathbf{y}$, corresponding fitted values $\hat{\mathbf{y}}$, and a vector of fitted weights which we shall denote by $\boldsymbol{\theta}$, this can be decomposed into a loss function and a penalisation term as follows

$$L(\mathbf{y}, \hat{\mathbf{y}}; \theta) = J(\mathbf{y}, \hat{\mathbf{y}}; \boldsymbol{\theta}) + R(\boldsymbol{\theta}) \quad (2)$$

The loss function $J(\mathbf{y}, \hat{\mathbf{y}}; \boldsymbol{\theta})$ in (2) depends on the ANN one is fitting, while $R(\boldsymbol{\theta})$ can take the form of the L^2 parameter norm penalty

$$R(\boldsymbol{\theta}) = \lambda \sum_{i=1}^{p} \theta_i^2$$

(which is the one we shall use, however other possible penalisation terms are the LASSO penalty and dropout regularisation). Since the ANNs we will be fitting are of the binary classification or regression type - for the former we take

$$J(\mathbf{y}, \hat{\mathbf{y}}; \boldsymbol{\theta}) = -\sum_{i=1}^{n} y_i \, log \, \hat{y}_i - \sum_{i=1}^{n} (1 - y_i) \, log(1 - \hat{y}_i)$$

while for the latter we take

$$J(\mathbf{y}, \hat{\mathbf{y}}; \boldsymbol{\theta}) = \frac{1}{N} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

Note that the y_i's and the $\hat{y}_i$'s are the components of $\mathbf{y}$ and $\hat{\mathbf{y}}$ respectively, and that for the binary classification problem the $\hat{y}_i$'s take the form of probabilities.

The approach we shall take is that of applying ANNs in two stages, and through the combination of their outputs, predict the policy loss for each policy. The first neural network will be trained on all the data to predict the probability of a claim. The second neural network will be trained to predict log claim severity of the policy, given a non-zero occurrence - in this case, the predicted claim severity is then retrieved by back-transforming using the exponential function. We consider the log claim when fitting neural networks to allow negative outputs and avoid having our neural network severely affected by very large claims. For any set of inputs, we finally obtain the predicted policy loss by multiplying the predicted claim probability for said inputs with the predicted claim severity for same inputs if a claim occurs.

DATA AND MODELLING APPROACHES

Obtained from an insurance company database after signing a non-disclosure agreement, the dataset we use contains 214840 motor insurance policies, of which 15441 are claimant. The dates of the policies span a 9 year period from 2009 till 2018. The dataset contains a number of variables relating to both the policy holder as well as their vehicle, which due to non-disclosure terms we are not able to identify. We were also provided with the claim severity of each policy, which is non-zero in the case of a claimant's policy. It must be noted that there may be multiple observations pertaining to the same policy holders in separate years or for separate vehicles, but information identifying the individual is withheld, and therefore the independence assumption on observations will be made.

Preliminary Analysis: Distribution Fitting

Two exponential family distributions, the gamma and inverse Gaussian distributions, and the lognormal and Weibull distributions, are first fitted to claim severity data. The AIC and BIC are given in Table 1. It can be seen that the lognormal and the inverse Gaussian distribution have a considerably better fit than the gamma and Weibull distribution. The distribution fits for claim severity data overlapping the histogram can be seen in Figure 1. The good fit shown by the inverse Gaussian distribution to the claim severity indicates the suitability of the ZAIG model to the application.

Table 1: AIC and BIC of each of the Proposed Distributions

Proposed Distribution	AIC	BIC
Gamma	258290	258305
Weibull	258831	258846
Lognormal	255329	255344
Inverse Gaussian	255409	255424

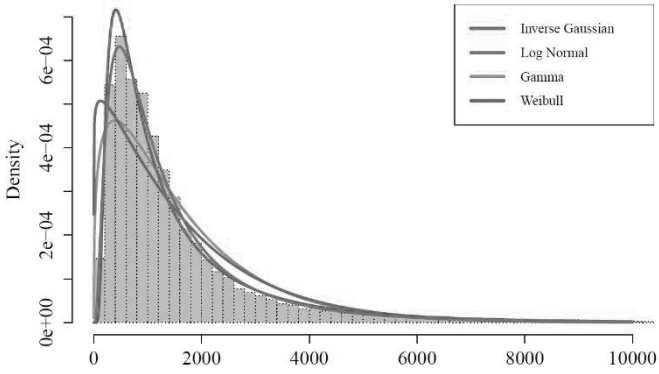


Figure 1: Histogram of Claim Severity Overlapped by the Probability Density Functions of Various Theoretical Distributions (for Clarity Purposes, the x-axis of this Histogram has been Trimmed to Display Claims of Severity of up to €10000)

The next step involves the modelling of claim probability, claim severity and policy loss given a number of inputs. Due to the data's sensitive nature, besides for the aforementioned response variables, we shall not disclose the specific variable labels of the inputs within the dataset and their nature, i.e. whether covariate or factor. The aim is to make use of both the ZAIG model and the two-stage neural network approach to predict claim occurrence, claim severity and policy loss.

Fitting the ZAIG Model

Several ZAIG models are fitted by making use of a package called *gamlss* in R. Literature suggests that when the number of observations is large, as in this case, the model with the best AIC is chosen, as the heavy parameter penalty on the BIC will cause the model to be too simplistic. Let ph_{ij} refer to aforementioned ZAIG-type terms involving the j^{th} explanatory variable concerning the i^{th} policy holder's

characteristics, while let car_{ij} refer to the j^{th} explanatory variables concerning the i^{th} policy holder's car. Furthermore, let $const_{\mu}$, $const_{\sigma}$ and $const_{\pi}$ be intercept terms. The model with the smallest AIC of 360874.1 is given by the equations:

$$\log(\mu_i) = const_{\mu} + \sum_{j=1}^{3} ph_{ij} + \sum_{j=1}^{3} car_{ij}$$
$$\log(\sigma_i) = const_{\sigma} + car_{i1}$$
$$\text{logit}(\pi_i) = const_{\pi} + \sum_{j=1}^{4} ph_{ij} + \sum_{j=1}^{4} car_{ij}$$

As mentioned earlier, the inputs deemed influential will shall not be disclosed due to the sensitive nature of the data. We now move on to describing the implementation of the two-stage ANN approach.

Implementing the Two-stage Artificial Neural Network Approach

All the ANN modelling carried out is coded in Python with a *Keras* back-end, and the models are run on cloud computers provided by Google Colaboratory for research purposes. When working with ANNs, there are a number of decisions that must be made by the practitioner prior to training the model. These include deciding upon the following hyperparameters (i) the number of hidden layers of the neural network; (ii) the number of neurons in each hidden layer (we take this to be common for all layers); (iii) the activation functions used between layers; (iv) the learning rate of the SGD algorithm; and (v) the batch size of the SGD algorithm. For both cases, all combinations of the following settings are tried: (i) hidden layers: 1, 2, 3; (ii) neurons in hidden layers: 8, 16, 32, 64; (iii) activation function: ReLU, Leaky ReLU; (iv) learning rate: 0.01, 0.001, 0.0001; and (v) batch size: 16, 32, 64 (for the claim probability model) and 4, 8, 6 (for the claim severity model).

We first construct the ANN model for the claims probability. We find the optimal neural network models by attempting all the above combinations and evaluating the fit of each combination by means of 3-fold cross-validation. *K*-fold cross-validation is a popular cross-validation technique which divides the learning set randomly into *K* partitions, and each partition is used as a validation set for a model fitted on the remaining learning set. The *K* validation measures on each partition are then averaged to produce a single measure. In this case, we calculate the mean of the loss of each model over the three cross-validation folds. This can be seen in Table 2. The standard deviation of the loss indicates it would be naive to single out one of the top 5 models as the overall best model. Therefore, with the aim of reducing the possibility of overfitting, the final claim probability ANN model will be constructed through the use of the basic ensemble method (BEM) by averaging the output of the top 5 models (Perrone 1993). Similarly, the ANN model for claims severity is constructed, where the activation function of the output layer is taken to be the identity function and the output is taken to be the logarithm of the claim. Table 3 yields the results. Once again, the BEM is used to average the output of the top 5 models. As mentioned earlier, for any given input, the predicted policy loss is obtained by multiplying the outputs of the two ANN-based models. In the next section, we compare the two different approaches to assess in which aspects one approach is superior to the other.

Table 2: Top 5 Performing Claim Probability ANN Models from the Grid Search. Mean and Standard Deviation of the Loss Function Presented in the Table.

Model	mean (loss)	s. d. (loss)
(1,64,Leaky ReLU,0.01,32)	0.25810	0.00235
(1,32,Leaky ReLU,0.01,32)	0.25813	0.00227
(1,16,Leaky ReLU,0.01,32)	0.25815	0.00225
(1,16,Leaky ReLU,0.01,64)	0.25825	0.00236
(1,64,Leaky ReLU,0.01,64)	0.25827	0.00230

Table 3: Top 5 Performing Claim Severity ANN Models from the Grid Search. Mean and Standard Deviation of the Loss Function Presented in the Table.

Model	mean (loss)	s. d. (loss)
(3,64-32-16,Leaky ReLU,0.001,4)	-0.8845	0.0149
(2,64-16,Leaky ReLU,0.001,8)	-0.8854	0.0127
(3,64-64-32,Leaky ReLU,0.001,4)	-0.8861	0.0134
(2,32-32,Leaky ReLU,0.001,8)	-0.8862	0.013
(3,32-32-16,Leaky ReLU,0.001,8)	-0.8864	0.0137

COMPARATIVE ANALYSIS

For both approaches, we first compare the performance for predicting claim probability and claim severity separately. The first is done on the whole dataset while the second is done on the claimants' policies. Finally, we compare the performance of both approaches in predicting the overall claim on all the dataset. In all cases, we will make use of 5-fold cross-validation to test the performance of both modelling approaches.

For claim probability models, we use two goodness-of-fit criteria. The first is the negative loglikelihood for the Bernoulli distribution, which is the negative of $J(\mathbf{y}, \hat{\mathbf{y}}; \boldsymbol{\theta})$ for the classification neural network. We can use this also on the ZAIG model by substituting $\hat{y}_i$'s with the predicted values for the π_i's. The second is the well-known normalised Gini coefficient. 5-fold cross-validation is performed and the results are presented in Table 4. It can be seen that the negative log likelihood of the classification ANN model is slightly smaller than that of the ZAIG model, while the normalised Gini coefficient is also larger. Both these results indicate that the classification ANN model is a marginal improvement over the ZAIG model.

Table 4: Negative Log Likelihood and Normalised Gini Coefficient Results for Claim Probability Models

Model	Neg. Log. Lik.	Normalised Gini
ZAIG	0.24574	0.35112
ANN	0.24417	0.37463

For claim severity models, we use three commonly known goodness-of-fit criteria: the root mean square error (RMSE), the mean absolute error (MAE) and the median absolute error (MedAE). The latter two, in particular MedAE, are more robust to outliers. 5-fold cross-validation is performed and the results are presented in Table 5. It is interesting to note that the ZAIG model results in a slightly lower value of the RMSE. On the other hand, the regression ANN yields a lower MAE and a considerably lower MedAE. This indicates that the ZAIG model predictions get closer to outliers while the ANN approach is less affected by outliers and better at predicting non-extreme claims.

Table 5: RMSE, MAE and MedAE Results of Claim Severity Models

Model	RMSE	MAE	MedAE
ZAIG	1935.6548	1134.0340	809.4621
ANN	2036.1180	1018.7798	526.4716

Lastly and most importantly, we wish to judge the models based on how well they are able to predict policy loss. We do this by comparing the predicted policy loss from the ZAIG model and the two-stage neural network approach to the actual loss. Using the usual 5-fold cross validation approach, we obtain the results in Table 6. It is interesting to note that while the RMSE is fairly similar for both models, with the ZAIG model faring slightly better than the two-stage ANN approach, the MAE and MedAE are significantly smaller for the latter. This indicates that the two-stage ANN approach is better than the ZAIG model at predicting the loss of non-extreme observations.

Table 6: RMSE, MAE and MedAE Results of Policy Loss Models

Model	RMSE	MAE	MedAE
ZAIG	660.7622	202.5629	88.2330
2-stage ANN	663.3498	168.6635	50.7894

CONCLUSION

In this study we compare the standard ZAIG model for predicting policy loss to a two-stage ANN approach, both of which we implement on a considerably large data set. We conclude that while both approaches are highly valid for this application, the two-stage ANN approach fares slightly better when predicting class probabilities, and significantly better when predicting claim severity and claim size of non-extreme observations. For overall observations, the two methods are comparable with the ZAIG model faring slightly better. As a result, the two-stage ANN approach has its advantage as it does not put the burden of extreme observations on the policy holders when adopting a pricing strategy based on these predictions. Indeed, in practice, extreme claims can also be catered for via reinsurance contracts. We therefore promote this approach as a viable alternative tool for pricing policies in insurance.

REFERENCES

Bortoluzzo, A.; D. Claro; M. Caetano; and R. Artes. 2011. "Estimating total claim size in the auto insurance industry: a comparison between Tweedie and zero-adjusted inverse Gaussian distribution." *Brazilian Administration Review* 8, 37-47.

Brockman, M.J. and T.S. Wright. 1993. "Statistical motor rating: making effective use of your data." *Journal of the Institute of Actuaries* 119, No.3, 457-543.

Dalkilic, T.E.; F. Tank; and K.S. Kula. 2009. "Neural networks approach for determining total claim amounts in insurance." *Insurance: Mathematics and Economics* 45, No.2, 236-241.

Glorot, G.; A. Bordes; and Y. Bengio. 2011. "Deep sparse rectifier neural networks." In *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics*, 315-323. Available online.

Goodfellow, I.; Y. Bengio; and A.C. Courville. 2015. *Deep Learning*. The MIT Press.

Gschlößl, S. and C. Czado. 2008. "Modelling count data with overdispersion and spatial effects." *Statistical papers* 49, 531-552.

Haberman, S. and A.E. Renshaw. 1996. ``Generalized linear models and actuarial science." *Journal of the Royal Statistical Society Series D (The Statistician)* 45, No.4, 407-436.

Heller, G.; D. Stasinopoulos;. and R. Rigby. 2006. "The zero-adjusted inverse Gaussian distribution as a model for insurance claims." In *Proceedings of the 21st International Workshop on Statistical Modelling*, 226-233. Available online.

Jorgensen, B. and M.C. Paes De Souza. 1994. "Fitting Tweedie's compound Poisson model to insurance claims data." *Scandinavian Actuarial Journal* 1994-1, 69-93.

Maas, A.L.; A.Y. Hannun and N.Y. Ng. 2013. "Rectifier nonlinearities improve neural network acoustic models." In *Proceedings of International Conference on Machine Learning* 30, No.1, 3. Available online.

Noriszura, I. 2013. "Estimation of claim cost data using zero-adjusted gamma and inverse Gaussian regression models." *Journal of Mathematics and Statistics* 9, 185-192.

Perrone, M.P. 1993. "Improving regression estimation: averaging methods for variance reduction with extensions to general convex measure optimization." PhD Thesis, Brown University Providence, RI, USA.

Pesantez-Narvaez, J.; M. Guillen; and M. Alcan{\~n}iz. 2019. "Predicting motor insurance claims using telematics data - XGBoost versus logistic regression." *Risks* 7, No.2, 1-16.

Pinquet, J. 1997. "Allowance for cost of claims in bonus-malus systems." *ASTIN Bulletin, Cambridge University Press* 27, No.1, 33-57.

Rigby, R.A. and D.M. Stasinopoulos. 2005. "Generalized additive models for location, scale and shape." *Journal of the Royal Statistical Society Series C (Applied Statistics)* 54, No.3, 507-554.

Smyth, G.K. and B. Jorgensen. 2002. "Fitting Tweedie's compound Poisson model to insurance claims data: dispersion modelling." *ASTIN Bulletin, Cambridge University Press* 32, 143-157.

Ye, C.; L. Zhang; M. Han; Y. Yu; B. Zhao; and Y. Yang. 2018. "Combining predictions of auto insurance claims." arxiv.org/abs/1808.08982.

Yunos, Z. M.; A. Ali; S.M. Shamsuddin; I. Noriszura; and R. Sallehuddin. 2016. "Predictive modelling for motor insurance claims using artificial neural networks." *International Journal of Advances in Soft Computing and its Applications* 8, No.3, 160-172.

AUTHOR BIOGRAPHIES

LUKE AAROHI is a graduate in Mathematics, Statistics and Operations Research at the University of Malta, currently working as a trader in the Maltese financial industry.

DAVID SUDA has a Ph.D. in Statistics from Lancaster University and is currently a full-time lecturer and researcher at the University of Malta. He has also provided training and consultation to public and private entities in Malta, and is currently also participating in collaboration with other academics on a number research projects.

CLOUD BASED SIMULATION

Public Cloud Data Storage and Retrieval Security

Sonia Amamou[a] , Zied Trifa[b] and Maher Khmakhem[c]
[a] MIRACL Research Center, Sfax, Tunisia
[b]ISAE Gafsa, University of Gafsa, Tunisia
[c]University of King Abdulaziz, College of Computing and Information Technology

Keywords*:* Public Cloud , Data Protection , storage , Cloud Security , Security Model , Data storage , Data retrieval , retrieval

Abstract *:* Data protection is considered to be a main facet in order to have a favorable development of the infrastructure of public cloud. This is due to the fact that it mainly indicates the well working of infrastructure especially for the cloud users and also for the cloud providers. Unfortunately, data protection has not yet been provided with the required attention. Nowadays, in order to achieve data protection, cryptographic approaches are recommended as solutions because they can grant some degree of control on data users. Nevertheless, they can reduce the efficiency and also set up an important overhead of any cloud system . In this work, we studied the storage and retrieval in the cloud, which are protection methods. Then, we evaluated their strengths and also their limitations in order to propose a new approach.

1. Introduction

The rise of the Internet and the democratization of computing have mostly led to the emergence of cloud computing with McCarthy1 and Arpanet2 around the 1960s. However, its true trigger came into existence in 2002 when Amazon opted for renting its IT resources not in use; out of holiday season, with third-party organizations .The cloud concept was, then, incorporated with Salesforce3. In fact, this proposed a customer relationship management (CRM) as a service. Consequently, Cloud was considered by IBM as a real true business. Recently, Cloud has become a favorable trend for any company that maintains the services it offers and its benefits such as execution speed, cost reduction, computing power, etc. This market significance is transfigured into numbers. In other words, the researches carried out by some organizations such as Sales force and Medinsoft4 suppose that: "By 2020, the global cloud market is expected to exceed $ 241 billion. 840 million people around the world will use cloud computing solutions, saving businesses a minimum of $ 210 billion a year "(Saleforce and Medinsoft, 2012). Moreover, the cloud has granted a significant gain in time and also energy:" Still, the company must deeply think about security issues, such as protecting its data, because in case of violation it will have an outstanding impact on its brand image, its strategies as well as the confidence of its shareholders ,before moving to the cloud.

2. Cloud definition

Cloud computing provides opportunities to have online services accessible at any time for any client. Precisely, the cloud allows each user to make a simple request via any website or to share cloud offers with a provider, a solution or infrastructures. As for e-commerce, cloud computing is one of the most vague technique terminologies in history. One reason is that cloud computing can be used in many application scenarios. Furthermore, cloud computing are intensively promoted by various companies for business promotion. Additionally, we can see that the cloud computing is in the fast growing phase when referring to Hyper Cycle published by Gartner Group in 2008 [1] [2].

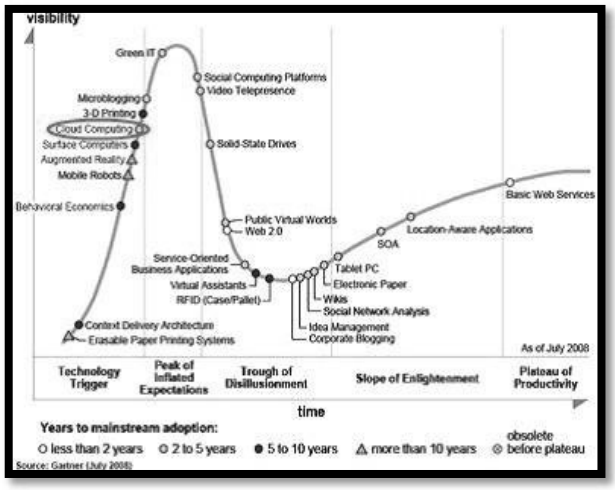


Figure 1 : Cloud computing is in the phase of fast growing [3]

3. Cloud computing architecture

A number of organizations and researchers have established the architecture for cloud computing. The whole system can be divided, essentially, into core stack and management.Firstly, there are three layers in the core stock which are (1) Resource (2) Platform and (3) Application.

In fact, the resource layer is the infrastructure layer that comprises physical and virtualized computing, storage as well as networking resources.

Then, the platform layer is the most complex. This is due to its division into many sublayers.E.g. a computing framework manages the transaction dispatching and/or task scheduling. A storage sub-layer supplies unlimited storage and caching capability.

Finally, the application server and some other components hold up the same general application logic as before with either on-demand capability or flexible management, such as no element will be the bottle neck of the whole system [2].

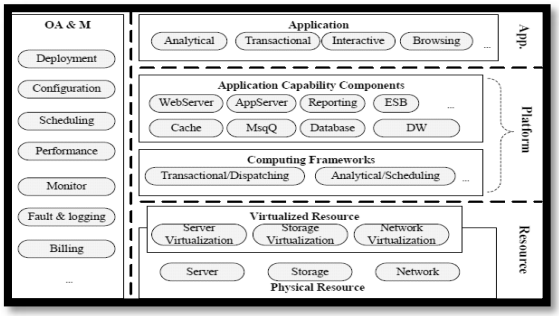


Figure 2 : The Reference architecture [3]

With reference to the underlying resource and the components, the application could support substantial and distributed transactions. Moreover, it can support also management of voluminous data. This is to say that all the layers contribute external service along web service or other open interfaces.

4. Cloud computing categories

As a matter of fact, there is a lot of dimensions that allow us classifying cloud computing. The two most common frequent used categories areservice type and service boundary.

- From the service type's view, cloud computing can be classified as Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). SaaS lays out services in order to end users, while IaaS and PaaS provide services to ISV anddevelopers leaving a margin for 3-party application developers.
- From the service boundary's view, cloud computing can be divided into public cloud, private cloud and hybrid cloud. First and foremost, the public cloud refers to services provided to external parties. Secondly, the private cloud operates for enterprises; They establish it by themselves. Finally, Hybrid cloud shares resources between public cloud and private cloud using a secured network such as Virtual Private Cloud (VPC) services released by Google[4] and Amazon [5].

4.1 Cloud Deployment Types

Clouds have main deployment types (also referred to as deployment models in some references):

- Public Cloud: The Cloud supplier owns the infrastructure of a public Cloud which is, in turn, hired to Cloud clients. The Cloud supplier systematically supervises its physical framework, but it can acquire particular functions to a third party as in the case of procuring hardware maintenance such as in Amazon and RackSpace .
- Private Cloud: A private Cloud deployment sort is preserved and used by a particular enterprise which means that workers are its only clients. Accurately, the private Cloud could either oversee its own infrastructure or obtain the management to a third party. This includes most infrastructures of banks and telecommunication companies.
- Community Cloud: Organizations sharing common business functions and/or objectives could cooperate and found their own particular community such as Associated Newspaperswhich is a group of newspapers and publishing media that build up a community Cloud framework to serve their common goals.
- Hybrid Cloud. This deployment combines private, community, and/or public Cloud which is vital to reinforce higher adaptability, accessibility, and also reliability [6].

5. Methods used to secure data transfer:

5.1. Anonymization

The anonymization [7] of data involves that the information directly identifying individuals will be deleted or generalized. In fact, it includes 2 categories:

- **Generalization** which consists of diluting attributes by adjusting their scale to make them less precise. It also associates the k-anonymity and its variants l-diversity and t-proximity
- **Randomization** weakens the link between

the attribute and the individual by adding noise to less rely the information. It involves noise addition, permutation and differentialprivacy.

Nevertheless, anonymization is subject to certain risks that can cause the re-identification of individuals.

- **Individualization:** possibility of isolating an individual in a database.
- **Correlation:** connecting two or more records relative to the same target individual.
- **Inference:** possibility to conclude the value of an attribute associated to an individual.

5.2 K-Anonymity

Around the 1990s, an experiment was conducted by a US researcher to show that anonymity is not enough to intercept people from being re-identified. As a matter of fact, he used a database published by a health facility and supposed to be anonymous but with "postal code / sex / date of birth» triplet kept.
By linking each existing record in the database with the convenient triplet appearing in a public list of electors, he has managed to recognize 80% of people [8].

5.3 L-Diversity

The model of l-diversity [9] expands k-anonymity by adding an additional constraint on the equivalence classes: each attribute must have L different values so that the sensitive fields become less precise and definite.

6. Conclusion

In short, cloud computing is a dematerialized IT model allowing its users to outsource their data and resources to a third-party supplier. This shift to a new environment can harrass the customer because of several factors such as trust to the provider, data geo-location, resources collocation and also security. In truth, several methods have been suggested in order to protect the data being transferred to the cloud, including encryption, anonymization and tokenization. There is also another complementary method which is monitoring.
In this context and based on the works done, we have tried to propose a model addressing safety issues based on encryption, anonymization and monitoring through sampling.

References

[1] Dimitrios Zissis and Dimitrios Lekkas. Addressing Cloud Computing Security Issues. In Proceedings of the international journal of Future Generation Computer System, vol.3, pp.583–592, 2012.

[2] Jinqwei Li; Dan Lin; Anna Squicciarini; Jin Li; Chunfu Jia. Towards Privacy-Preserving Storage and Retrieval in Multiple Clouds. IEEE Transactions on Cloud Computing, 1 October 2015 .

[3] Ling Qian, Zhiguo Luo, Yujian Du, and Leitao Guo , Cloud Computing: An Overview , January 2009 .

[4] http://appengine.google.com

[5] Amazon Web Service, http://aws.amazon.com

[6] Dawoud, W., Takouna, I., and Meinel, C. Infrastructure as a Service Security: Challenges and Solutions. In Proceedings of the 7th International Conference on Informatics and Systems, pp. 1–8, 2010.

[7] G. Murthy and R. Srinivas. Achieving Multidimensional k-Anonymity by a Greedy Approach. In Proceedings of the International Conference on Web Services Computing, 2011.

[8] Mayil Vel Kumar, P., Karthikeyan, M. L-diversity on k-anonymity with external database for improving privacy preserving data publishing. In Proceedings of the International Journal of Computer Applications, vol. 54, No.14, 2012.

[9] Sweeney, L(2002). Achieving k-Anonymity Privacy Protection Using Generalization and Suppression. Carnegie Mellon University, International Journal on Uncertainty, Fuzziness and Knowledge-based Systems, Vol n10, p 571-588.

R-TNCES VERIFICATION: DISTRIBUTED STATE SPACE ANALYSIS PERFORMED IN A CLOUD-BASED ARCHITECTURE

Chams Eddine Choucha
Faculty of Mathematical, Physical and Natural Sciences,
University of Tunis-El Manar, Tunisia.
Email: chamseddinechoucha@gmail.com

Naima Souad Ougouti
LSSD Laboratory, Computer Science Department,
University of science and technology of Oran Mohamed Boudiaf, Algeria.
Email: souad.ougouti@univ-usto.dz

Mohamed Khalgui
National Institute of Applied Sciences and Technology,
University of Carthage, Tunisia.
Email: khalgui.mohamed@gmail.com

Laid Kahloul
LINFI Laboratory, Computer Science Department,
Biskra University, Algeria.
Email: laid.k.b@gmail.com

KEYWORDS
R-TNCES, Formal verification, Computation tree logic CTL, Cloud Computing, Distributed architecture

Abstract

This paper deals with formal verification of reconfigurable discrete event control systems (RDECSs) using reconfigurable timed net condition/event systems (R-TNCESs) formalism. A reconfigurable system switches from a mode to another during its working process to adapt its behavior to the related environment. By including such a feature, RDECSs become complex and their verification is often expensive in terms of computation time and memory. R-TNCES verification consists of two major steps: state space generation and state space analysis. In order to improve the analysis step, we proposed in this paper, a new approach for CTL properties verification performed on a developed distributed cloud-based architecture. The proposed approach allows to avoid redundant calculation and to reduce time execution by considering relationships between properties and setting up a parallelization tree that ensures a suitable execution order for each property. The evaluation of the proposed approach is achieved by exploiting a case study, which illustrates the impact of using this approach. The results show the relevance of this paper.

INTRODUCTION

Reconfigurable discrete event control systems (RDECSs) Zhang et al. (2018) are new complex systems, which can be manufacturing systems Khalgui and Hanisch (2011), real time systems and intelligent control systems Khalgui et al. (2011). RDECSs satisfy several conditions (concurrency, control, communication, etc). In fact, RECESs are the trend of future systems. However, the complexity of RDECSs makes formal verification of the latter complex. Model checking remains the most used formal technique for its automatic verification, it consists of checking functional properties satisfaction in order to validate the behavior of a formal model. Recently, industry and academia have shown a strong interest in the formal verification of reconfigurable systems Hafidi et al. (2018). Many researchers developed new formalisms to deal with reconfigurable systems and proposed new works on improving formal verification methods. Reconfigurable net condition/event systems (R-TNCESs) are developed as an extension of TNCES formalism in Zhang et al. (2013), where reconfiguration and time properties with modular specification are provided in the same formalism. Authors in Camilli et al. (2014) devloped a CTL Model checker in cloud using map reduce. The basic idea is to increase computation power and data availability to reduce time execution. They perform distributed fixed-point algorithm. However, they do not consider similarities that system model may contain, which implies redundant calculation during verification. Both of layer by layer verification proposed in Zhang et al. (2013) and the formal verification proposed in Hafidi et al. (2018) focused on the improvement of state space generation, thus, they neglect state space analysis. In this paper, we are interested in reconfigurable discrete event control systems, modeled by the R-TNCES formalism where RDECS behavior is represented by the behavior of control components (CCs) and the communication

between them Zhang et al. (2013). Indeed, we propose a new verification method that aims to optimize the R-TNCESs formal verification. State space analysis is a complex step of verification, thus we focus on its improvement. The proposed approach consists to deploy an adequate distributed architecture to perform parallel verification for CTL properties. This distributed architecture is a hybrid cloud-based architecture composed of: 1. A local workstation, which performs simple tasks on a unique computation machine. 2. A virtual workstation, which performs complex tasks on the cloud using several virtual machines. The objective of this work is to reduce both execution time and number of CTL properties to be verified. In this aim, we first proceed to CTL properties classification using methods proposed in Ramdani et al. (2018), which allow us to distinguish between simple and composed CTL properties and to define possible relationships between them. The results of classification are then used by an algorithm to compute a matrix relationships. This algorithm allows us to extract different relationships that models can contain. Thus, it allows us to detect redundant properties. A second algorithm is used to generate a parallelization tree. Indeed, by exploring the tree using BFS algorithm, we set a suitable order for each property to be checked. Finally, we will proceed to the verification of CTL properties on the allocated virtual machines. Each virtual machine uses SESA tool Zhang et al. (2013) to perform verification and stores the result in the shared memory. We can summarize our work as follows 1. Development of a hybrid distributed cloud-based architecture, where we perform parallel CTL properties verification. 2. CTL properties classification in order to check if a property is simple or composed and to detect possible relationships. 3. Matrix relationships computation, where information about relationships between each couple of properties are mentioned. 4. parallelization tree generation, which allows us the guiding of verification in order to avoid redundant calculation. To demonstrate the feasibility our approach, a formal case study is provided. We detail the application of each verification process step to show how the proposed method works from the classification to the CTL properties verification. Moreover, we perform some simulations as a preprocessing before the model-checking process in order to illustrate the gain of this approach. The results show that the proposed technique reduces effectively the execution time and the number of properties to be verified for reconfigurable systems verification. The organization of this paper is as follows. Section 2 gives backgrounds. Section 3 presents the formalization. The papers contribution is applied to a formal case study in Section 4. Finally, Section 5 concludes this paper and describes future works.

BACKGROUND

In this section, we present required concepts to follow the rest of the paper.

Reconfigurable Timed Net Condition/Event System

R-TNCES is an extension of TNCES Hanisch et al. (1997). It is defined in Zhang et al. (2013) as a couple $RTN = (B, R)$, where R is the control module consisting of a set of reconfiguration functions $R = \{r_1, ..., r_n\}$ and B is the behavior module, which is a union of multi TNCES-based CC modules, represented by

$$B = (P; T; F; W; CN; EN; DC; V; Z_0) \qquad (1)$$

where, 1. P (resp, T) is a superset of places (resp, transitions), 2. $F \subseteq (P \times T) \cup (T \times P)$ is a superset of flow arcs, 3. W: $(P \times T) \cup (T \times P) \rightarrow \{0, 1\}$ maps a weight to a flow arc, $W(x, y) > 0$ if $(x, y) \in F$, and $W(x, y) = 0$ otherwise, where $x, y \in P \cup T$, 4. $CN \subseteq (P \times T)$ (resp, $EN \subseteq (T \times T)$) is a superset of condition signals (resp, event signals), 5. DC : $F \cap (P \times T) \rightarrow \{[l_1, h_1], .., [l_{F\cap(P\times T)}, h_{F\cap(P\times T)}]\}$ is a superset of time constraints on input arcs of transitions, where $\forall$ i$\in$ [1, $|F \cap (P \times T)|$], l_i, h_i $\in \mathbb{N}$ and $l_i < h_i$, 6. $V : T \rightarrow \wedge, \vee$ maps an event-processing mode (AND or OR) for every transition; 7. $Z_0 = (M_0, D_0)$, where $M_0 : P \rightarrow \{0, 1\}$ is the initial marking, and $D_0 : P \rightarrow \{0\}$ is the initial clock position.

Computation Tree Logic

Computational tree logic CTL is a temporal logic for branching time. CTL models system evolution as a tree-like structure where each state can evolve in several ways (i.e., specify behavior systems from an assigned state in which the formula is evaluated by taking paths) Camilli et al. 2014. In CTL formulas, we use temporal operator *next* (X) and *until* (U) to express that a property is valid in the next state, and until another property becomes valid. The operators *finally* (F) expresses that a property will eventually become valid, while *globally* (G) expresses that a property is valid along the entire subsequent path. However ,in CTL temporal operators must be immediately preceded by The universal path operator (A) or the existential path operator (E) to express that a property is valid for all and for some paths, respectively. If AP is the set of atomic propositions, and $p \in AP$, CTL formulas are defined as follows:

$$\phi ::= p|\neg\phi|\phi \wedge \phi|A\psi|E\psi \text{ (state formulas).}$$
$$\psi ::= X\psi|F\phi|G\phi|\phi U\psi \text{ (path formulas).}$$

Relationships

Each reconfiguration of an RDECS is represented by an TNCES, where each physical process is represented by a control component CC. The execution

order creates relationships between physical processes that express dominance and equivalence.

Definition 1. A Dominance relationship is created when the execution respects the order of precedence between physical processes.

Definition 2. An equivalent behavior is created when the execution includes the same physical process.

Relationships between physical processes reflected on CTL properties. Thus, every two properties that share a goal or a part of goal can be in a be-part relationships Ramdani et al. (2018). It can be: (i) a dominance relationship when the final result of the whole formula depends only on the dominant part of this formula because of the dominance relationship between the physical process. (ii) equivalent relationship. We denote two types of properties: simple goal properties and composed goal proprieties.

Definition 3. Simple goal properties have a simple goal represented by a simple place or simple transition.

Definition 4. Composed goal proprieties have a composed goal represented by a conjunction of places or transitions or liveness properties, they may be expressed with CTL using the operator presented in Ramdani et al. (2018).

Cloud Computing

The NIST Mell et al. (2011) defines the cloud computing as a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be provisioned and released with minimal management effort Hayes (2008). With the rise of the cloud computing trend, several providers have emerged, including the giants Amazon, Google and Microsoft that propose different storage and computing services. The proposed services allow to deploy secure and visualized architectures on-demand and in a large scale via internet.

Motivation

The apparition on new complex systems makes RDECS verification task complex. Real RDECSs encompass millions of transitions, which increase considerably time verification and number of properties to be verified. We propose in this paper a new approach for CTL properties verification. Thus, we aim to make model checking more efficient by reducing both time validation and number of properties to be verified. The improvement of RDECSs verification is obtained by considering different possible relationships, which allows us to propose an optimal scheduling for the CTL properties verification . Moreover we propose to perform a parallel properties verification on a distributed architecture deployed in a cloud computing server. Figure 1 shows how we proceed to parallel CTL properties verification.

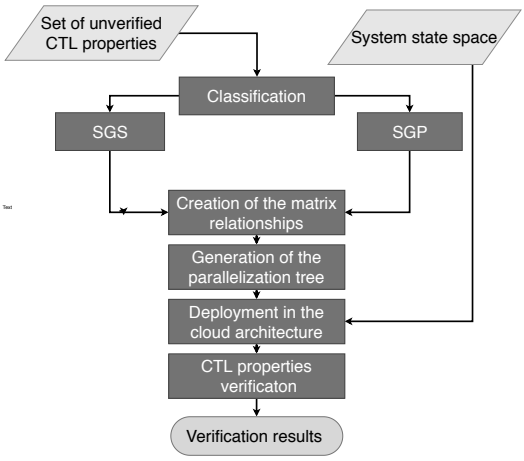


Figure 1: Parallel verification of CTL properties.

Formalization

In this section, we present CTL properties verification on a distributed cloud-based architecture. Our approach can be summarized by the following steps.
Step 1. Relationship detection: We adopt algorithms developed in Ramdani et al. (2018) to detect different relationships between TNCESs. We denote by

- Dom^{C_i} (resp. Dom^{C_i,C_j}) operator that express dominance relationships between two formula on the same configuration (resp. dominance relationships between two formula on two configurations).

- $Equi^{C_i,C_j}$ operator that express equivalence relationships between two formula on two configurations.

Step 2. Matrix relationships: From the result of the precedent step, a square matrix of size n is created, where n is the number of properties to be verified. Each element of the matrix define if there is or not a relationships between a couple of formulas. $M[i,j] = 1$ means that P_i *Dom* P_j, $M[i,j] = 2$ means that P_i is equivalent to P_j and $M[i,j] = 0$ means that P_i is not related to P_j.
Step 3. Parallelization tree generation: Using algorithm 1 and matrix relationships MR, we generate parallelization tree in order to coordinate the execution of properties verification, thus to avoid redundancy. Classically, we have to check all properties in blind way, but thanks to this step we identify the redundancies and the factorization between them.
Step 4. CTL properties verification: CTL properties are affected according to the established execution oder to the allocated virtual machines VMs, where each VM proceed to CTL properties verification using SESA tool.

Distributed cloud-based architecture for CTL properties verification

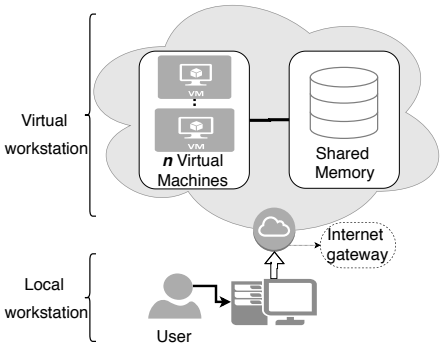


Figure 2: Hybrid distributed cloud-based architecture for CTL properties verification.

In this subsubsection, we present the developed hybrid distributed cloud-based architecture. Figure 2 shows that architecture core contains two distinguishable parts as follows. 1. A local workstation, which is a unique computational machine to perform simple computational tasks (Grouping operation, matrix relationships computation and parallelization tree generation). 2. A virtual workstation (i.e., on the cloud), which is a virtual private network that perform complex computation tasks composed of: A data storage unit, which is a shared memory to store state spaces, MRs, RTs and the result of each property verification. n virtual machines VMs, with $n = Min(10, Max(NOfNode(RT))$. Where, (i) $NOfNode(RT)$ is the function that takes a parallelization tree RT and returns the number of nodes present in each level. (ii) $Max()$ resp. (Min) returns the maximum (resp. minimum) among a set of natural numbers.

Implementation

In this subsection, we present the main algorithms exploited in our approach.
Matrix relationships: The used algorithm creates a square matrix relationships of size s the number of properties to be verified. each element of the matrix describes the relationship between two properties P_i and P_j. $M[i,j] = 0$ when there is no relationship between P_i and P_j, $M[i,j] = 1$ when P_i Dom P_j and $M[i,j] = 2$ when P_i $equi$ P_j.
Parallelization tree generation: Algorithm 1 generates a tree, where properties relationships are represented by a parent-child relation. ex. The relationship $P_2 Dom^{C_1} P_6$ is represented by two node P_2 and P_6, where P_2 is the parent and P_6 the child.

Algorithm 1 Parallelization tree generation

```
Input: M[Card(P_CTL), Card(P_CTL)]: Matrix;
Output: A: Tree;
Begin
Create_root(); CurrentNode ← root;
for i = 1..Size(M) do
 if (Marked(P_i)) then
    Create_Chiled(P_i); Mark(P_i);
    for j = 1..Size(M) do
       if (M[i,j]==1) then
          Create_Chiled(P_i,P_j); Mark(P_j);
       end
       if (M[i,j]==2) then
          Mark(P_j;)
       end
    end
    else
       for j = 1..Size(M) do
          if (M[i,j]==1) then
             Create_Chiled(P_i,P_j); Mark(P_j);
          end
          if (M[i,j]==2) then
             Mark(P_j);
          end
       end
    end
 end
end
return A;
end
```

EXPERIMENTATION

Case Study

In this subsection, we demonstrate the relevance and the gain of the proposed contribution. We apply the proposed approach on an R-TNCES $RTN_s = \{R_{s_1}, B_{s_1}\}$ that model a sequential system s_1. The system s_1 is composed of eight physical process. Figure 3 shows the module behavior B_{s_1}. RTN_{s_1} is able to be reconfigured between three configuration (C_1,C_2, C_3) due to a user requirement or environment changes. The configurations are described by the following control chains.
$C_{Chain_1} = (CC_1, CC_2, CC_3, CC_4)$.

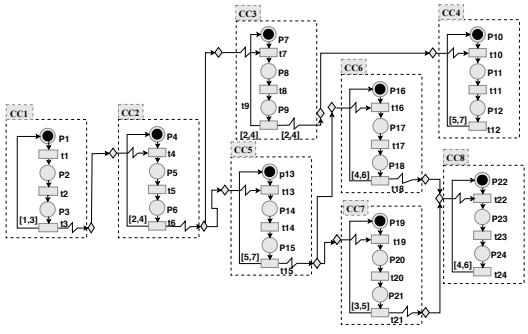


Figure 3: Model with three configurations process.

$C_{Chain_2} = (CC_1, CC_2, CC_3, CC_5, CC_6, CC_8)$.
$C_{Chain_3} = (CC_1, CC_2, CC_3, CC_5, CC_7, CC_8)$.
According to control module, RTN_{s_1} can apply six reconfiguration as follows. $R_{s_1} = \{(C_1, C_2); (C_1, C_3); (C_2, C_1); (C_2, C_3); (C_3, C_1); (C_3, C_2)\}$.

Application

Table 1

Set of CTL properties ϕ	
SGP	$P_1 : EF(p_3)$, $P_2 : AF(p_4)$, $P_3 : EF(p_6)$, $P_4 : EF(p_9)$, $P_5 : EF(p_{12})$, $P_6 : EF(p_{15})$, $P_7 : EF(p_{18})$, $P_8 : EF(p_{21})$, $P_9 : EF(p_{21})$, $P_{10} : EF(p_{24})$.
CGP	$P_{11} : AG(p_{15} \rightarrow EF(p_{21}))$. $P_{14} : AG(p_3 \rightarrow EF(p_{24}))$. $P_{13} : AG(p_{15} \rightarrow EF(p_{18}))$. $P_{12} : AG(p_9 \rightarrow EF(p_{12}))$.

System configurations may behave incorrectly. In order ton ensure the correctness of system modeling, we must verify that reconfigurations fulfill systems properties. We denote the set of required properties on the behavior by ϕ. Where, $\phi = \{P_i\}_{i=1,..,14}$. First, We apply tasks that are performed in the local workstation. Classification and grouping operations: We adopt algorithms proposed in Ramdani et al. (2018). Table 1 presents CTL properties in ϕ and the results of grouping operations (i.e., shows if $P_i \in SGP$ or $P_i \in CGP$). Tabe 2 shows the relationships between CTL properties in ϕ.

Table 2: Relationships between CTL properties.

Relationships between properties in ϕ	
Equivalence	$P_8 Equi^{C_3} P_9$
Internal dominance	$P_6 Dom^{C_2} P_{13}$
External dominance	$P_4 Dom^{C_2 C_3} P_{12}$. $P_6 Dom^{C_2 C_3} P_{11}$. $P_1 Dom^{C_2} P_{14}$.

Then, we perform matrix relationships algorithm. The created matrix is used to perform parallelization tree generation algorithm 1 to generate the tree RT shown in Figure 4.

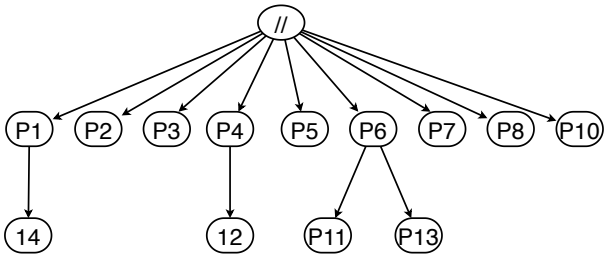


Figure 4: Parallelization tree.

Virtual machine allocation: We allocate n virtual machine, with $n = Min(10, Max(NOfNode(RTN_{s_1})$. Thus, $n = Min(10, 9)$. In a second step, we perform properties verification on virtual workstation. VMs use SESA tool to proceed to $P_i, i = 1..14 \in \phi$ verification. In order to define execution order of properties verification, we explore the parallelization tree RT using BFS algorithm. In the present case, execution order is define as follows: $Level1 : P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8, P_{10}$. $Level2 : P_{12}, P_{11}, P_{13}, P_{14}$.

Evaluation

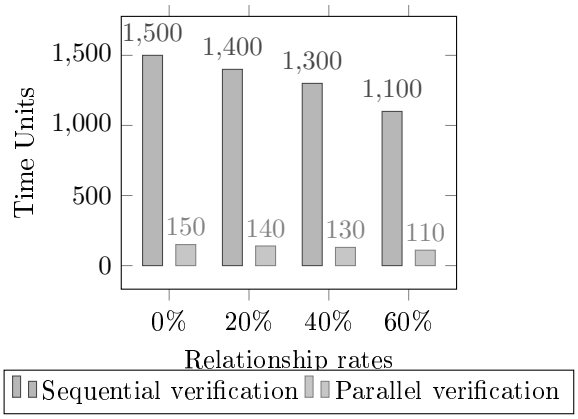


Figure 5: Sequential verification VS Parallel verification.

Let assume we have to verify a system model with 250 TNCESs. In order to ensure the well-behave of the system we have to verify at least 4 properties for each TNCES. Thus we need to verify 1000 CTL properties. Relationship rates can be: (i) Low in $0, 20\%$, (ii) $Medium$ in $20, 60\%$, or (iii) $High$ when more than 60%. Figure 5 presents the result of properties verification with relationships consideration in a parallel and sequential way. The results shows that the gain increases proportionally to relationships rate. Thus, the gain is clearly shown when relationships rate is $High$. Moreover, we can observe an important gain when performing parallel verification thanks to the proposed architecture. The proposed architecture allows us to reduce considerably times execution by (i) Avoiding redundant calculation, (ii) Avoiding wait time execution, (iii) establishing an execution order for properties verification. The number of properties to be verified increase proportionally with system size and number of its TNCECs. In order to ensure TNCES correctness, we must check at least 4 properties. Due to our proposed architecture, we aim to deal with large scale system verifications. Let assume a set of systems that have 200, 500 and 1000 TNCESs, with $Medium$ rate 40% of dominance relationships and 10% of equivalence relationships. Figure 6 shows that our proposed approach is more efficient when dealing with huge systems. Figure 7 compares the results of the proposed approach with the approach proposed in Zhang et al. (2013), it shows that proposed one allows to reduce execution time thanks to the parallelization and relationships consideration.

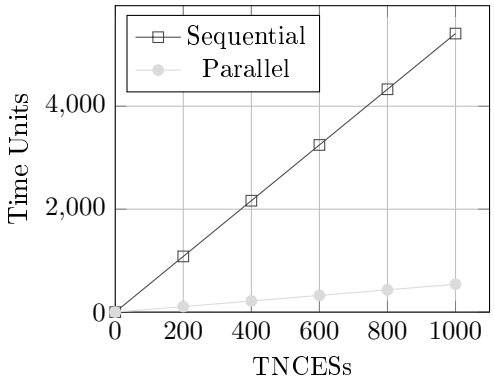


Figure 6: Improved performance of Parallel verification.

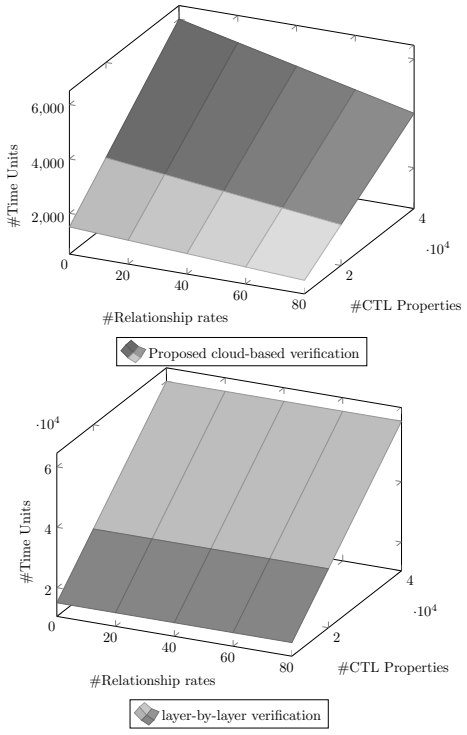


Figure 7: Proposed approach vs layer by layer verification.

Conclusion

This paper deals with formal verification of RDECSs modeled with R-TNCES. The proposed method aims to improve state space analysis by using a hybrid distributed cloud-based architecture for computation tasks. Developed architecture is composed of: 1. A local workstation, where simple computation tasks are executed. First, a classification algorithm is applied in order to distinguish between simple and composed properties. Then, we compute a matrix relationships that mention any eventual relationship between each couple of properties. Finally, we generate a parallelization tree that we explore to extract a suitable execution order for each property to be verified. 2. A virtual workstation, where complex tasks are computed. Virtual machines use SESA tool to perform CTL properties verification and stores results in the shared memory. The proposed contributions allow us to increase computational power, data availability and to perform parallel execution in order to reduce time execution and the number of properties to be verified. This work opens several perspectives; first, we plan to apply our approach in verification of real-case studies and to develop an extension of this method for dealing with reconfigurable distributed behaviors. Second, we plan to perform a distributed state space generation and analysis using a cloud-based architecture.

References

Camilli M.; Bellettini; et al., 2014. *CTL model checking in the cloud using MapReduce.* In *Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), 2014 16th International Symposium on.* IEEE, 333–340.

Hafidi Y.; Kahloul L.; Khalgui M.; Li Z.; Alnowibet K.; and Qu T., 2018. *On Methodology for the Verification of Reconfigurable Timed Net Condition/Event Systems. IEEE Transactions on Systems, Man, and Cybernetics: Systems*, , no. 99, 1–15.

Hanisch H.M.; Thieme; et al., 1997. *Modeling of PLC behavior by means of timed net condition/event systems.* In *Emerging Technologies and Factory Automation Proceedings, 1997. ETFA'97., 1997 6th International Conference on.* IEEE, 391–396.

Hayes B., 2008. *Cloud computing. Communications of the ACM*, 51, no. 7, 9–11.

Khalgui M. and Hanisch H.M., 2011. *Reconfiguration protocol for multi-agent control software architectures. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 41, no. 1, 70–80.

Khalgui M.; Mosbahi O.; Li Z.; and Hanisch H.M., 2011. *Reconfiguration of distributed embedded-control systems. IEEE/ASME Transactions on Mechatronics*, 16, no. 4, 684–694.

Mell P.; Grance T.; et al., 2011. *The NIST definition of cloud computing.*

Ramdani M. et al., 2018. *Automatic Properties Classification Approach for Guiding the Verification of Complex Reconfigurable Systems.* In *Proceedings of the 13th International Conference on Software Technologies, ICSOFT 2018, Porto, Portugal, July 26-28, 2018.* 625–632.

Zhang J.; Frey G.; Al-Ahmari; et al., 2018. *Analysis and control of dynamic reconfiguration processes of manufacturing systems. IEEE Access*, 6, 28028–28040.

Zhang J.; Khalgui M.; Li Z.; Mosbahi O.; and Al-Ahmari A.M., 2013. *R-TNCES: a novel formalism for reconfigurable discrete event control systems. IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 43, no. 4, 757–772.

INTERNET OF THINGS

Dynamic Anomaly Detection based on Recursive Independent Component Analysis of Multi-Variate Residual Signals

Edwin Lughofer
Department of Knowledge-Based Mathematical Systems, Johannes Kepler University Linz
email: edwin.lughofer@jku.at

Mahardhika Pratama
School of Computer Science and Engineering, Nanyang Technological University
email: mpratama@ntu.edu.sg

Christian Eitzinger
Profactor GmbH, Steyr-Gleink, Austria
email: christian.eitzinger@profactor.at

Thomas Radauer
Stratec Consumables, Anif, Austria
email: t.radauer@stratec.com

KEYWORDS
anomaly detection in industrial facilities, multi-sensor measurements, multi-variate residual signals, causal relation networks, on-line update of causal relation models, recursive independent component analysis (RICA), on-line production system

ABSTRACT

We address the problem of anomaly detection in industrial facilities based on multi-sensor measurement systems. Any anomaly may point to an (early) indication of a fault in the system, which may turn latter into a severe failure and thus should be detected as early as possible. Our approach relies on causal relation networks gained from the measurements of process variables, whose influence weights between causes and effects are modelled by a specific non-linear fuzzy systems architecture comprising generalized rules. Residuals are extracted online from a bunch of (high-quality) causal relations and are analyzed over time through an *advanced independent component analysis (ICA)*, which employs a specific strategy on dominant parts for automatically suppressing noise content in the data/residuals. A further important novelty aspect is that the mixing and demixing matrices which characterize the independent components are recursively updated in incremental and robust (converging) manner (thus, our approach is termed as *RICA*). The control signal extracted from the components is therefore also dynamically updated. The second dynamic aspect concerns the update of the causal relation models themselves through the usage of an evolving fuzzy systems approach. This integrates the possibility of structural changes by decreasing or increasing the models' non-linearity degree on demand and on the fly. RICA is compared with static ICA and with several (15) related state-of-the-art (SoA) methods based on data sets from a chip manufacturing system embedding two production stages. They comprise OK process phases with different machining parameters and NOT-OK (abnormal) phases. Results show improved detection capabilities with RICA while achieving lower false alarms rates than static ICA and than the SoA methods.

INTRODUCTION

Motivation and State-of-the-Art

In todays large-scale manufacturing systems and Industry 4.0 processes, the supervision of the whole production chain for approaching *zero defect manufacturing* (Myklebust 2013) and realizing predictive maintenance (Lughofer and Sayed-Mouchaweh 2019) with high performance can be established through the usage of *data-driven anomaly and fault detection methods* (Chandola et al. 2009). Such methods have been successfully employed in IoT platforms within Industry 4.0, see the recent survey (Behniafar et al. 2018) or the framework proposed in (Chen et al. 2012). Thereby, an utmost goal is to detect any possible arising anomaly in the system at an early stage in order to take action to avoid subsequent faults (e.g., by job re-scheduling techniques (Nouiri et al. 2017) or by process optimization cycles (Lughofer et al. 2019)). The early recognition is important (and the focus of this paper) because faults may lead to severe machine breakdowns or often induce downtrends in the quality of the production (Lughofer et al. 2018b) which leads to defective production parts/items. This in turn induces significant waste for the company and/or even complaining customers (as in our use case, see also Section 'Application Scenario').

Compared to predictive and forecast modeling methods used for anomaly/fault prognostics (Malathi et al. 2010) (Chen 2013) (Djeziri et al. 2016), the essential point is that anomaly detection operates in a fully *unsupervised manner*. This circumstance triggers an applicability of anomaly detection also to a wide range of (on-

line) quality control systems, where quality, health or failure information, needed to establish proper prognostics and/or classification models, is costly to obtain or cannot be provided at all (Wang and Gao 2006).
Current anomaly detection methods range from pure statistical approaches in form of recognizing outliers from regular data distributions representing fault-free modes (Zhang et al. 2018) through measuring the deviations in transformed data (Odgaard et al. 2008) and (model) parameter spaces (Chen et al. 2014) to one-Class classification methods such as one-Class SVMs (Fernandez-Francos et al. 2013) or support vector data description (Duan et al. 2016). A survey of various methods can be found in (Chandola et al. 2009) (and 15 out of these will be compared with our new approach in the results section). These methods are operating directly in the original or transformed feature space, but do not provide any interpretation aspects of model outputs, neither in the model structures at all.
This situation has been improved in previous works such as (Serdio et al. 2017) or (Lughofer 2018), where partial SysID models have been trained from data to explain relations, dependencies contained in the system when everything is working properly. The training is conducted based on a multi-variate analysis which determines the input-output structure purely on a correlation basis (without looking for actual cause-effect relations). Residuals are extracted for new on-line samples, characterizing the degree of deviations to the SysID models. These works, however, neglect the analysis of the multi-variate residuals directly in their joint space, as they solely act on a uni-variate basis and produce fault warnings by a simple OR-operator. Thus, these approaches are not able to characterize more complex (hidden) structures among residual signals which may become more clearly or even earlier 'violated' in the case of anomalies than in separate single control signals (especially, when *intermittent faults* occur (Luo 2006)).

Our Approach

Our approach overcomes these bottlenecks by i.) looking for real (=statistically dependent) causal relations in the system based on a modified version of the PC algorithm (Spirtes et al. 2000) as a preliminary filter step, ii.) building up non-linear regression models based on the causal network structure with the usage of generalized TS fuzzy systems (Lughofer et al. 2015) — for each causal relation in the network, a non-linear (fuzzy) model is established for characterizing the (local) degree of influence of each cause onto a particular effect —, and iii.) by a multi-variate residual signal analysis, realized through an advanced independent component analysis (ICA). The latter is able to characterize hidden structures in the multi-dimensional residual space through latent variables (LVs). Significant changes in the characteristics of the LVs (energy content of reconstruction), subject to an automated control limit, then indicate anomalies. Thereby, noise suppression (leading to a reduction of false alarms) is elegantly embedded by operating solely on the dominant parts of the demixing matrix. A further novelty aspect of our approach is that it takes into account the (possible) dynamic nature of processes, which may lead to deteriorations in the models and thus in the residual signal quality.
To achieve this, we propose the following two extensions:

1) A recursive update of the demixing and mixing matrices with new on-line samples, which are used to reconstruct the components (whose energy content is used as control signal) based on new residual vectors. The recursive update of these is achieved in a way such that it allows single-pass, sample-wise updates, and that no additional parameters are required. The recursive derivation assures convergence to the (hypothetical) batch solution.
2) An update of the (partial) causal relation models contained in the network. This update integrates a recursive parameter adaptation method converging to the classical weighted least squares solution and a structural evolution component. The latter allows the addition or removal of rules on the fly in order to increase or decrease the non-linearity degree of the models ($\rightarrow$ adaptive non-linearity). Thereby, the structure of the network remains unchanged in order to meet the stability aspect from relational point of view.
A summary of our framework is visualized in Figure 1, where the upper part denotes the (off-line) system identification step and the lower part represents the analysis of the multi-variate residual signals during on-line execution. Therein, the two components shown as thick boxes with bold font entries denote the two new issues regarding model adaptation and update of independent components which will be discussed below.
Our approach was evaluated and successfully compared with various related SoA methods, including the univariate residual-based methods in (Serdio et al. 2017), see Table 2. This was done within the scope of a microfluidic chip production system. Therefore, particular data sets (from on-line measurements) have been drawn from the process, including the regular process behavior under anomaly-free conditions and this also with changing machining parameters (to check for robustness with respect to these). They also include a real problematic case where customer complaints about the quality of the chips arose. The comparison showed improved detection capability as well as lower false alarm rates than the SoA methods, and this with much lower parametrization effort.

ESTABLISHING THE FUZZY-REGRESSION CAUSAL RELATION NETWORK

Our approach is based on two stages: i.) seeking for causal relations in the fully spanned variable space to explain the internal real statistically significant cause-

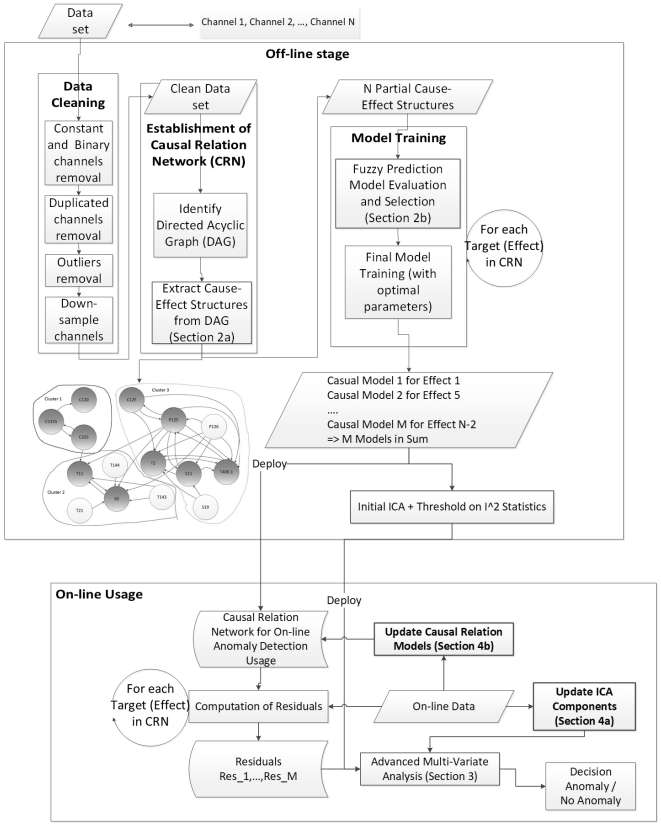


Figure 1: Framework of the anomaly detection strategy; sections where components are described in more detail enumerated as such; the dynamic components for updating models and independent components are shown in bold font.

effect structures between the system variables; ii.) using the found partial causal structures for establishing multi-variate (non-linear) regression models. Thereby, each effect in the CRN is used as target and each cause for each effect as input. The regression coefficients for each cause then indicate the impact/weight of this cause on the effect, in relation to the other causes, rather than showing the conditional dependencies between causes and effects (as achieved in Bayesian networks (Jensen 2001)).

The first stage is realized through the usage of a fast version of the PC algorithm (Spirtes et al. 2000) (Le et al. 2014), which operates in two steps. In the first step, it learns from data a skeleton graph, which contains only undirected edges. In the second step, it orients the undirected edges to form equivalence classes of directed acyclic graphs (DAGs). Each equivalence class can then be interpreted as a cause-effect relation, where the nodes denote the variables (sensor channels) and the directed edges define the cause-effects. For the theoretical foundations of the PC algorithm, please refer to (Spirtes et al. 2000).

In the second stage, the concrete causal relation models between all causes and effects are established. Therefore, each variable appearing as effect in the CRN, is used as a target variable and each cause flowing into the same effect is used as an input variable. We aim for both, i) robustness with respect to noise and ii) multivariate characteristics of the relation model. The combination of both can be best achieved with a regularized multi-variate regression approach, where the regression coefficients directly represent the relative weights and a regularizer in the objective function guarantees for robustness in the case of noise and high-dimensional input spaces. Furthermore, the causal relation may have a different characteristics with different levels of influences between causes and effects in different parts of the feature space. Thus, a local partitioning of the regression model into different sub-models may be an important issue as well. Therefore, we employ the structure of Takagi-Sugeno (TS) fuzzy systems, which are able to decompose the input space into partial local regions (rules) and to represent local relations in the context of (weighted) multi-variate regression. Functionally, a TS fuzzy system is defined as:

$$\hat{f}(\vec{x}) = \sum_{i=1}^{C} \Psi_i(\vec{x}) \cdot l_i(\vec{x}) \quad \Psi_i(\vec{x}) = \frac{\mu_i(\vec{x})}{\sum_{j=1}^{C} \mu_j(\vec{x})} \quad , \quad (1)$$

with C the number of rules (local partitions) and $l_i(\vec{x}) = w_{i0} + w_{i1}x_1 + ... + w_{ip}x_p$ a linear hyper-plane with input dimensionality p. In the generalized case, the membership degree $\mu_i(\vec{x})$ to the ith rule is estimated by

$$\mu_i(\vec{x}) = \exp(-\frac{1}{2}(\vec{x} - \vec{c}_i)^T \Sigma_i^{-1}(\vec{x} - \vec{c}_i)) \quad (2)$$

with $\vec{c}_i$ the center and Σ_i^{-1} the inverse covariance matrix of the ith rule, describing its shape and orientation.

For learning the non-linear antecedent parameters and the adequate number of rules, we interpret the data set as a pseudo-stream and apply *Gen-Smart-EFS* learning engine Lughofer et al. (2015), which is a single-pass method embedding incremental merging and splitting operations Lughofer et al. (2018a) for establishing homogenous and compact rules. This procedure extract the appropriate number of rules in a single pass, where the rule evolution criterion decides when to evolve new rules. It is based on funded statistical quantile theory exploring the usage of the prediction interval as a tolerance region of a rule. Positioning and shapes of the rules are determined synchronously and recursively by *Gen-Smart-EFS*.

Afterwards, the following regularized objective function is minimized to find the regression coefficients $\vec{w}_i$ of each local linear model $l_i(\vec{x}), i = 1, ..., C$ separately:

$$J_i = \sum_{k=1}^{N} \Psi_i(\vec{x}(k)) e_i^2(k) + \lambda \sum_{j=1}^{p} (\alpha w_{ij}^2 + (1-\alpha)|w_{ij}|)$$

where $e_i(k) = y(k) - \hat{y}_i(k)$ represents the error of the local linear model in the kth sample (N samples in sum), λ the regularization parameter and α a parameter in $[0, 1]$. In combination, both terms are also used in the elastic net approach Zou and Hastie (2005) (for classical global linear regression). Thus, by solving the weighted objective J_i through using cyclical coordinate descent method along regularization paths Hastie et al. (2010), and this with weights given by the normalized rule membership degrees $\Psi_.$, we term our approach as *fuzzily weighted elastic net*. This weighted optimization of consequent parameters per rule separately has several advantages over a joint global one regarding robustness and interpretability — see Lughofer (2011) (Chapter 2).

ADVANCED RESIDUAL SIGNAL ANALYSIS

Once all partial relations embedded in the CRNs have been established, it is possible to check new incoming on-line samples how well they 'fit' into the network. They can be checked by inspecting the residuals between real (y_i's) and predicted ($\hat{y}_i$)'s effects:

$$res_i = y_i - \hat{y}_i \quad \forall i = 1, ..., M \tag{3}$$

M is the number of causal relations found and used for on-line processing due to a sufficient quality of the regression fit. Checking each residual signal $res_i(k)$ (k denoting the kth sample/time instance) separately (as done in (Serdio et al. 2017)), has bottlenecks in terms of i.) univariate residual analysis with narrow views (no synergies / dependencies among the residuals are checked), ii.) fixed widths of tolerance bands being tunable only in a posteriori manner (through ROC curves) and iii.) manually designed filters requiring significant parametrization effort for noise suppression. Thus, in order to get rid of these shortcomings, we propose the extraction of a control signal with fully automated control limits from the energy content of independent components of the *joint multivariate residuals*, and this with automated noise suppression.

Control Limits by Independent Component Analysis

Independent component analysis (ICA) is a technique which attempts to find latent variables (LVs) characterizing the joint residual space in a statistically independent manner (Comon 1994). In our setting of multivariate residual signals R, the relation between the original signals and the latent variables is thus described as:

$$R = A * S + E \tag{4}$$

where $R \in \mathbb{R}^{M \times N}$ is the residual matrix composed of N residual vectors comprising residuals from M models, which are assumed to reflect a kind of mixture or superpositions of (hidden) independent (latent) variables stored in $S \in \mathbb{R}^{D \times N}$. This assumption is realistic because usually several variables may appear as either causes or effects in different partial causal relations. $A \in \mathbb{R}^{M \times D}$ (with D the reduced dimensionality of LVs) denotes the mixing matrix, containing the basic information about the hidden dependence structure contained in the multi-variate residuals. Therefore, its counterpart, the demixing matrix $W \in \mathbb{R}^{D \times M}$ contains the basic information about the *hidden independence structure*, i.e. the knowledge how to demix a multivariate residual to achieve independent components S. Finding A and W can be achieved in iterative optimization cycles — as carried out in the *fast ICA* method (Hyvärinen and Oja 2000).

The idea for the purpose of anomaly detection is now to use the demixing matrix W (determined from the training data) in order to demix new residual vectors $\vec{r}_{new} \in M \times 1$ into the (nearly) independent LVs $\vec{s}_{new}$: $\vec{s}_{new} = W\vec{r}_{new}$. Whenever $\vec{s}_{new}$ shows an atypical appearance compared to the vectors $\vec{s}_1, ..., \vec{s}_N$ obtained on the training data, it may point to a change in the hidden independence structure. Such an atypical appearance can be recognized due to a significant increase of the energy of the s-vectors, which can be measured in terms of squared L^2-norm of the vectors (termed as I^2 statistics), thus by:

$$I^2(new) = \vec{s}_{new}^T \vec{s}_{new} \tag{5}$$

We address the noise problematic by not using the whole matrix W, but only the most *dominant parts* in W. Thereby, our idea is to re-arrange the rows of W in descending order according to their L^2 norms $\rightarrow W_b$ and to take the first X% stored in W_d when multiplying with $\vec{r}_{new}$, thus

$$\vec{s}_{new} = W_d * \vec{r}_{new} \tag{6}$$

This means that we omit non-dominant parts (with low absolute weights in the LVs) for the reconstruction step of a new residual vector. The non-dominant parts are typically more intensively affected by noise in the data than the dominant parts. In this sense, we suppress or at least dampen atypical noise occurrences in new online test data without needing an explicit design of filters requiring tuning parameters (as in (Serdio et al. 2017)), while assuming that the dominant parts are sensitive to real anomalies. For achieving appropriate control limits on $I^2(new)$, we adopt the idea of non-parametric kernel density estimation (KDE) for estimating the cumulative distribution function (CDF) of the I^2-statistics as achieved when applying it to the (anomaly-free) training data samples — see (Wand and Jones 1994) for details. Once having the CDF estimated, it is easy to extract an anomaly detection threshold as a confidence limit $1 - \alpha$, due to a significance level α for denoting an abnormal I^2 energy (typical values are 0.05, 0.025 and 0.01 — we used $\alpha = 0.01$ in *all* our experiments).

ON-LINE UPDATE OF INDEPENDENT COMPONENTS AND RELATION MODELS

In the case of time varying systems where dynamic regular changes occur, which are not part of any anomal system behavior, but which are more intended (see also Lughofer (2018)), the off-line trained (causal relation) models as well as the statistical constructs (for independent components representation) may become outdated over time. This leads to a successively increasing deterioration of their predictive performance and representations, which in turn typically decreases their anomaly detection capabilities.

In order to circumvent this, we propose incremental update algorithms for both, the mixing and demixing matrices as well as for the causal relation models. They can operate in single-pass, sample-wise manner, meaning, that, once a new sample is loaded, the matrices/models can be immediately updated based solely on this single sample. In particular, for the mixing matrix A, we consider recursive estimates in the form:

$$A(n+1) = \alpha_n A(n) + \beta_n \vec{r}(n)\vec{f}^T(n) \tag{7}$$

which is due to the possibility to represent the mixing matrix A as $E(\vec{r}(n)\vec{f}^t(n))$ Comon (1994), with E the expectation, $\vec{r}(n) \in M \times 1$ the residual vector obtained from the current (the nth) sample and $\vec{f}^T(n) = \Phi(W\vec{r}(n)) \in 1 \times M$, with Φ a component-wise non-linearity function. According to Hyvärinen et al. (2001) (Chapter 9), for super-Gaussian independent components, the component-wise nonlinearity is given by: $\Phi = -2tanh(W\vec{r}(n))$, whereas for sub-Gaussian independent components, the component-wise nonlinearity is given by $\Phi = tanh(W\vec{r}(n)) - W\vec{r}(n)$. To decide whether components are super-Gaussians or sub-Gaussians, the kurtosis trick over past projected samples $s(k) = W(k)\vec{r}(k), k = 1, ..., n-1$ is applied (see Meyer-Baese and Schmid (2014) (Chapter 8)). $\beta_n = 1 - \alpha_n = \lambda$, where α_n denotes the forgetting factor: lower values trigger a faster forgetting of older information contained in $A(n)$ and thus emphasizing new information contained in the current residual $\vec{r}(n)$. By direct application of the famous Sherman-Morrison formula to matrix pseudo-inversion of the recursion in (7), the following sample-wise update for the demixing matrix can be obtained:

$$W(n+1) = \tag{8}$$
$$\frac{1}{\alpha_n}\left(W(n) - \frac{\beta_n}{\alpha_n + \beta_n \vec{f}^T(n)W(n)\vec{r}(n)} W(n)\vec{r}(n)g^T(n)\right)$$

with $g(n) = W^T(n)\vec{f}(n)$. This update together with (7) is termed *RICA* (*recursive independent component analysis*).

The update of the causal relation models concentrates on the update of the non-linear and linear parameters contained in the non-linear fuzzy regression models. This means that no structural change in the network is pursued, because deeper changes in the relations usually always indicate several problems (thus anomalies) and especially run a bit against the physical interpretation spirit to explain valid dependencies in the system. The update of the fuzzy regression models is achieved through the *Gen-Smart-EFS* learning engine Lughofer et al. (2015), where the rule centers are updated by a modified version of vector quantization (with decreasing learning gain) and the inverse covariance matrices (describing the shapes of the rules) are updated by a recursive exact formula, analytically derived with the help of the Neumann series Lughofer and Sayed-Mouchaweh (2015). The consequent parameters are updated through recursive fuzzily weighted least squares Lughofer (2011), where the weights are represented by the rule membership degrees: in this way, local learning for each rule separately is achieved, increasing stability and the update speed Lughofer (2011). Rules are evolved based on a statistical tolerance region for multivariate Gaussians leading to the χ^2 quantiles estimate. Whenever the following condition holds:

$$\left(min_{i=1,...,C} \;\; mahal_i = \sqrt{(\vec{x} - \vec{c}_i)^T \Sigma_i^{-1} (\vec{x} - \vec{c}_i)}\right) > r_{win},$$
$$win = argmin_{i=1,...,C}(mahal_i)$$
$$r_{win} = \text{fac} * p^{1/\sqrt{2}} \frac{1.0}{(1 - 1/(k_{win}+1))^m} \tag{9}$$

with $\vec{c}_i$ the center of the ith rule and Σ_i^{-1} its inverse covariance matrix, a new rule is evolved. k_{win} denotes the support of the winth rule (i.e. the closest one) and m a factor, default set to 4.

APPLICATION SCENARIO AND EXPERIMENTAL SETUP

In our case study, we deal with the inspection of microfluidic chips used for sample preparation in DNA (deoxyribonucleic acid) sequencing. Currently, the chips are checked by means of image inspection in an a posteriori manner, where bad chips are sorted out once they have already been produced, based on machine learning classifiers. This, however, typically does not prevent unnecessary waste. Therefore, the idea in this research was to supervise the process data which are directly recorded at the production time of chips and which may already reflect atypical occurrences, much before they can be (manually) realized by (a posteriori) inspection.

Data Characteristics

In fact, our test case contains process data which were recorded during the production of chips with unsatisfactory functionality. This opened the question whether any anomalies in the process data could be realized, which occurred during the production of those chips. If so, appropriate actions could be taken to omit waste

Table 1: Characteristics of the applied data sets drawn from injection molding (IE03) and bonding (BL02).

Data Set	# of PVs	# of Samp.	Setting	Time Frame
O. **OK**, IE03	64	2974	AI	5 days
O. **Not-OK**, IE03	64	5360	AI	5 days
O. **OK**, BL02	15	3881	AB	5 days
O. **Not-OK**, BL02	15	2440	AB	3 days
3 O. **OK**, BL02	15	6548	BB	2/8/1 days

and to reduce costs. Therefore, we traced back the corresponding chip order in our data base and extracted the data having occurred during the production of this order. This data thus served as test data whether our approach is able to detect anomalies properly. Furthermore, we extracted process data occurring during the production of another chip order, which were explicitly labelled as OK and which were produced with the *same machining parameters* as the chip order which were not OK. In this way, we can check whether our approach is actually able to detect the anomaly during production and not a possible intended drift due to a change in the machining parameters. Furthermore, we also extracted another data set from a different production order, which was also OK, but which was produced with a *different machining parameter* setting. In this way, we can check how robust our approach is against changes in these parameters. We extracted the appropriate data from the two essential chip production stages, injection molding and bonding. Table 1 summarizes the characteristics of all data sets.

Experimental Setup for Training and Evaluation

The most sensitive training parameter in the fuzzy systems training procedure *Gen-Smart-EFS* (Lughofer et al. 2015), i.e. the factor in the statistical tolerance interval for deciding whether to evolve a new rule or not (see (9)), is optimized by a grid-search technique coupled with 10-fold cross-validation procedure. That factor value leading to the minimal cross-validation error is selected and a final TS fuzzy model is trained on the whole training data set. This is repeated for each partial causal relation in the network found by the PC algorithm. No parameter tuning was conducted for the online residual analysis phase: the update of the demixing matrix and CRNs is parameter-free, the factor fac in (9) was used from the off-line tuning phase. We performed an on-line analysis of residuals one time with static ICA, one time with recursive updates of the demixing matrix (RICA), one time with conventional SPE Statistics (results reported in the last three columns in Table 2). For comparison purposes with related SoA works, we applied the (15) unsupervised multi-dimensional methods as listed in Table 2 with parameter tuning effort as indicated in the last column.

RESULTS

Residual Analysis for OK And Not-OK Data Sets

7 causal relation models with significant qualities could be obtained. When inspecting several residual signals obtained from the partial causal relations independently in a univariate way, several distinct peaks exceeding the tolerance bands several times significantly could be observed, which resulted in many false alarms produced by the approach in (Serdio et al. 2017). This situation seemed to improve much when applying SPE and I^2 statistics, as visualized in the left plots in Figure 2. Still, when applying the SPE statistics (upper image) there occurred a distinct peak exceeding the threshold (shown as dashed horizontal line) significantly at the end of the anomaly-free test data (green lines). The left image in the second row shows the I^2 statistics on the dominant parts, which could the false alarm rate to 0. Furthermore, during the NOT-OK data set several consecutive samples (surrounded by ellipsoids) exceeded the control limit (shown as horizontal dashed line), thus the anomaly could be detected as significant with I^2 statistics. This could be even reinforced when permanently updating the demixing matrix by RICA, as shown in the lower left plot: the majority of consecutive samples (circumvented by the ellipsoid) violated the threshold. Similar results could be achieved on the injection moulding data, again with improved results by static ICA and especially RICA (compared to SPE), and this with even more discrimination intensity between abnormal and normal phases.

Residual Analysis for Additional OK Data

The three right plots in Figure 2 compare the SPE statistics (upper) and the adaptive and non-adaptive I^2 statistics on dominant parts for the additional OK data set when a different machining parameter setting has been used (third part in red). Obviously, the static I^2 statistics produced much lower false alarms (4 only in sum over a period of 6548 samples = 0.06%) than the SPE statistics (3.26%). Here it is remarkable that the adaptation of the demixing matrix together with the adaptation of the causal relation models brings a significant reduction of the energy content of the I^2 statistics (see the lower right image), which in turn avoids any false alarms. Together with the more intense violations in the case of the NOT-OK phase (lower left plot), it can be concluded that the adaptation of the demixing matrix and the causal relation models is able to increase the distinction between the NOT-OK phase and the OK phases with different parameter settings.

Comparison with Related Methods

Table 2 shows the comparison of the performance of the various related SoA methods used for evaluation purposes. Thereby, three measures are presented: i.) the detection capability ('Yes' or 'No'), ii.) the delay of detection, and iii.) the false alarm rate in %, i.e. the per-

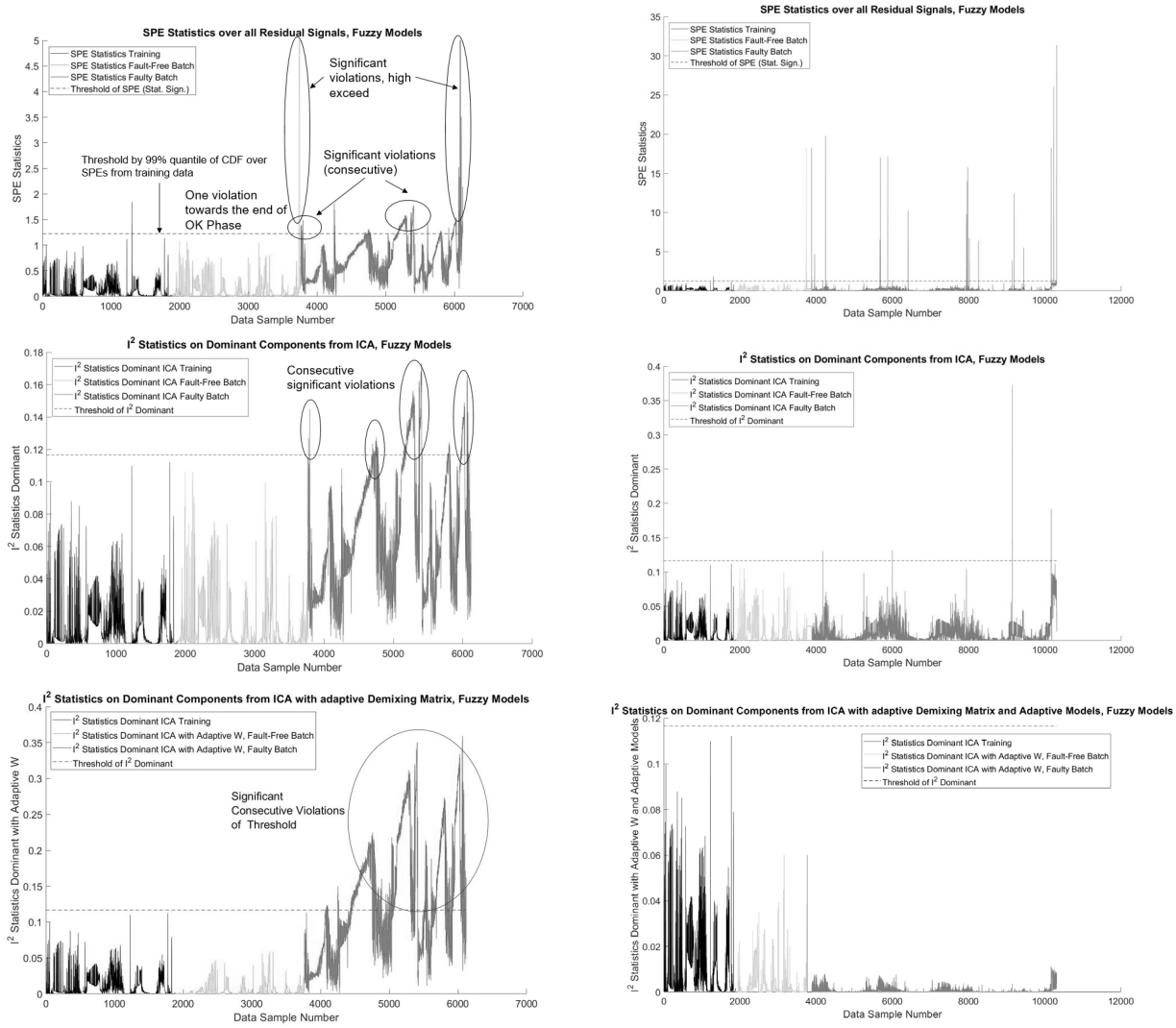


Figure 2: upper: SPE statistics over training data sets (black lines), test data sets of OK orders (greens) and test data set of NOT-OK orders / additional OK order (red lines); horizontal dashed lines indicate the control limits; left: BL02 with NOT-OK phase, right: BL02 with OK phase but different parameter setting; middle: the same for I^2 statistics on non-adaptive components; lower: the same for the I^2 statistics based on recursively updated components.

centage of samples wrongly classified as anomalies. Additionally, in the last column the parametrization effort of the methods is reported in terms of 'Low', 'Medium' and 'High, or even 'None'. The latter is the case for our approach, as the full tuning takes place during the off-line model building process.

From Table 2, it becomes clear that our new approaches based on I^2 statistics (reported in the last two rows) could outperform various related (unsupervised) SoA methods for anomaly detection as well as the previous uni-variate residual-based method in (Serdio et al. 2017), in terms of detection capability as well as false alarm rates. This is especially true for the bonding liner data set, where almost all methods (except support vector data description (SVDD) Duan et al. (2016) as an extension of one-class SVMs (Fernandez-Francos et al. 2013)) failed to detect the anomal behavior at all. In the case of injection molding (right part of the table), most methods can in fact detect the anomaly, but their false alarm rates (last but one column) are significantly higher than our variants. These higher false alarm rates are also the case for the bonding liner, except PCA with either Q-statistics or T^2-statistics (Odgaard et al. 2008) which produces a rate of 0 — however, these methods cannot detect the anomaly at all. The only exception in the case of bonding liner is SVDD, which can detect the anomaly with a slightly lower delay (1 vs. 3 samples), however, has a higher false alarm rate (1.51% vs. 0.05% with static ICA vs. 0.00% with RICA) and a much higher parametrization effort. Comparing the I^2 statistics of the static ICA based on static causal relation models with our RICA approach in combination

Table 2: Performance comparison of our (R)ICA-based control limits (last two rows) with 15 related unsupervised SoA methods; SPE is a direct measure of the energy content over multi-variate residuals, thus used for comparison (third to last row).

Methods	Bonding Liner			Injection Molding			Parametrization
	Det. Cap.	Delay	FA Rate (%)	Det. Cap.	Delay	FA Rate (%)	Effort
PCA, Q-Statistics	No	∞	0	Yes	25	1.78	None-Low
PCA, T^2-Statistics	No	∞	0	Yes	159	1.14	None-Low
Probabilistic PCA, Q-Statistics	No	∞	3.81	No	∞	1.28	None-Low
Probabilistic PCA, T^2-Statistics	No	∞	0	Yes	25	1.64	None-Low
kernel PCA, Q-Statistics	No	∞	3.44	Yes	1	1.7	None-Medium
kernel PCA, T^2-Statistic	No	∞	3.18	Yes	1	1.7	None-Medium
one-Class SVMs Linear	No	∞	2.76	Yes	1	9.58	High
one-Class SVMs Gaussian	No	∞	23.38	Yes	1	47.76	High
one-Class SVMs Sigmoid	No	∞	2.55	Yes	1	100	High
SVDD, Linear	No	∞	2.78	Yes	1	1.14	High
SVDD Gaussian	No	∞	8.62	Yes	1	48.97	High
SVDD SVMs Sigmoid	**Yes**	**1**	1.51	Yes	1	1.0	High
FD-kNN (Zhang et al. 2018)	No	∞	4.52	Yes	1	5.18	Medium
FD-wkNN (Zhang et al. 2018)	No	∞	4.57	Yes	1	3.34	Medium
Uni-variate Res. Analysis (Serdio et al. 2017)	Yes	1	4.12	Yes	1	0.92	Medium
SPE from CRNS	**Yes**	**2**	3.26	Yes	11	1.56	**None**
I^2 from ICA with CRNs (ours)	**Yes**	3	**0.05**	Yes	1	**0.43**	**None**
I^2 from RICA with adaptive CRNs (ours)	**Yes**	3	**0.0**	Yes	1	**0.22**	**None**

with adaptive CRNs (last but one versus last line), the latter can further decrease false alarm rates for both data sets (to 0% and to 0.22%). Furthermore, it is able to increase the intensity of correct detections in the case of bonding liner — see left plots in Figure 2. Decreasing FA rates towards 0% is essential in this production process to reduce efforts for wrong alarms and to avoid operator's mistrust in the anomaly detection system.

Acknowledgements

The first author acknowledges the support by the "LCM — K2 Center for Symbiotic Mechatronics" within the framework of the Austrian COMET-K2 program. The third and fourth authors acknowledge the Austrian research funding association (FFG) within the scope of the 'IKT of the future' programme, project 'Generating process feedback from heterogeneous data sources in quality control' (contract # 849962).

REFERENCES

Behniafar M.; Nowroozi A.; and Shahriari H., 2018. *A Survey of Anomaly Detection Approaches in Internet of Things. The ISC International Journal of Information Security*, 10, no. 2, 79–92.

Chandola V.; Banerjee A.; and Kumar V., 2009. *Anomaly Detection: A Survey. ACM Computing Surveys*, 41, no. 3.

Chen C.; Brown D.; Sconyers C.; Zhang B.; Vachtsevanos G.; and Orchard M.E., 2012. *An integrated architecture for fault diagnosis and failure prognosis of complex engineering systems. Expert Systems with Applications*, 39, no. 10, 9031–9040.

Chen H.; Tino P.; Yao X.; and Rodan A., 2014. *Learning in the Model Space for Fault Diagnosis. IEEE Transactions on Neural Networks and Learning Systems*, 25, no. 1, 124–136.

Chen M., 2013. *A hybrid ANFIS model for business failure prediction utilizing particle swarm optimization and subtractive clustering. Information Sciences*, 220, 180–195.

Comon P., 1994. *Independent Component Analysis: a new concept? Signal Processing*, 36, no. 3, 287–314.

Djeziri M.; Nguyen V.; Benmoussa S.; and Msirdi N., 2016. *Fault prognosis based on physical and stochastic models.* In *Proceedings of the 2016 European Control Conference*. 2269–2274.

Duan L.; Xie M.; Bai T.; and Wang J., 2016. *A new support vector data description method for machinery fault diagnosis with unbalanced datasets. Expert Systems with Applications*, 64, 239–246.

Fernandez-Francos D.; Martnez-Rego D.; Fontenla-Romero O.; and Alonso-Betanzos A., 2013. *Automatic bearing fault diagnosis based on one-class v-SVM. Computers & Industrial Engineering*, 64, no. 1, 357–365.

Hastie T.; Tibshirani R.; and Friedman J., 2010. *Regularized Paths for Generalized Linear Models via Coordinate Descent. Journal of Statistical Software*, 33, no. 1.

Hyvärinen A.; Karhunen J.; and Oja E., 2001. *Independent Component Analysis.* John Wiley and Sons Inc., New York.

Hyvärinen A. and Oja E., 2000. *Independent component analysis: algorithms and applications. Neural Networks*, 13, no. 4–5, 411–430.

Jensen F.V., 2001. *Bayesian Networks and Decision Graphs.* Springer Verlag, New York.

Le T.D.; Hoang T.; Li J.; Liu L.; Liu H.; and Hu S., 2014. *A fast PC algorithm for high dimensional causal discovery with multi-core PCs. Journal of LaTeX Class Files*, 13, no. 9, 1–13.

Lemos A., 2016. *Adaptive Fault Detection and Diagnosis using Evolving Intelligent Systems*. In *Proceedings of the IEEE Evolving and Adaptive Intelligent Systems Conference (EAIS) 2016*. Natal, Brazil.

Lughofer E., 2011. *Evolving Fuzzy Systems — Methodologies, Advanced Concepts and Applications*. Springer, Berlin Heidelberg.

Lughofer E., 2018. *Robust Data-Driven Fault Detection in Dynamic Process Environments Using Discrete Event Systems*. In M. Sayed-Mouchaweh (Ed.), *Diagnosability, Security and Safety of Hybrid Dynamic and Cyber-Physical Systems*, Springer, New York. 73–116.

Lughofer E.; Cernuda C.; Kindermann S.; and Pratama M., 2015. *Generalized Smart Evolving Fuzzy Systems. Evolving Systems*, 6, no. 4, 269–292.

Lughofer E.; Pratama M.; and Skrjanc I., 2018a. *Incremental Rule Splitting in Generalized Evolving Fuzzy Systems for Autonomous Drift Compensation. IEEE Transactions on Fuzzy Systems*, 26, no. 4, 1854–1865.

Lughofer E. and Sayed-Mouchaweh M., 2015. *Autonomous Data Stream Clustering implementing Incremental Split-and-Merge Techniques — Towards a Plug-and-Play Approach. Information Sciences*, 204, 54–79.

Lughofer E. and Sayed-Mouchaweh M., 2019. *Predictive Maintenance in Dynamic Systems — Advanced Methods, Decision Support Tools and Real-World Applications*. Springer, New York.

Lughofer E.; Zavoianu A.; Pollak R.; Pratama M.; Meyer-Heye P.; Zörrer H.; Eitzinger C.; and Radauer T., 2019. *Autonomous Supervision and Optimization of Product Quality in a Multi-Stage Manufacturing Process based on Self-Adaptive Prediction Models. Journal of Process Control*, 76, 27–45.

Lughofer E.; Zavoianu A.C.; Pollak R.; Pratama M.; Meyer-Heye P.; Zörrer H.; Eitzinger C.; Haim J.; and Radauer T., 2018b. *Self-adaptive evolving forecast models with incremental PLS space updating for on-line prediction of micro-fluidic chip quality. Engineering Applications of Artificial Intelligence*, 68, 131–151.

Luo M., 2006. *Data-Driven Fault Detection using Trending Analysis*. Ph.D. thesis, Department of Electrical and Computer Engineering, Louisiana State University and Agricultural and Mechanical College. Lousiana.

Malathi V.; Marimuthu N.S.; and Baskar S., 2010. *A comprehensive evaluation of multicategory classification methods for fault classification in series compensated transmission line. Neural Computing & Applications*, 19, 595–600.

Meyer-Baese A. and Schmid V., 2014. *Pattern Recognition and Signal Analysis in Medical Imaging*. Elsevier, Academic Press, Heidelberg.

Myklebust O., 2013. *Zero Defect Manufacturing: A Product and Plant Oriented Lifecycle Approach. Procedia CIRP*, 12, 246–251.

Nouiri M.; Bekrar A.; Jemai A.; Trentesaux D.; Ammari A.; and Niar S., 2017. *Two stage particle swarm optimization to solve the flexible job predictive scheduling problem considering possible machine breakdowns. Computers and Industrial Engineering*, 112, no. 12, 595–606.

Odgaard P.; Lin B.; and Jorgensen S., 2008. *Observer and Data-Driven-Model-Based Fault Detection in Power Plant Coal Mills. IEEE Transactions on Energy Conversion*, 23, no. 2, 659–668.

Serdio F.; Lughofer E.; Zavoianu A.C.; Pichler K.; Pichler M.; Buchegger T.; and Efendic H., 2017. *Improved Fault Detection employing Hybrid Memetic Fuzzy Modeling and Adaptive Filters. Applied Soft Computing*, 51, 60–82.

Spirtes P.; Glymour C.; and Scheines R., 2000. *Causation, Prediction, and Search, 2nd Edition*. MIT Press, Cambridge.

Wand M.P. and Jones M.C., 1994. *Kernel Smoothing*. CRC Press, Boca Raton.

Wang L. and Gao R., 2006. *Condition Monitoring and Control for Intelligent Manufacturing*. Springer Verlag, London, UK.

Zhang C.; Gao X.; Li Y.; and Feng L., 2018. *Fault Detection Strategy based on Weighted Distance of k Nearest Neighbors for Semiconductor Manufacturing Processes. IEEE Transactions on Seminconductor Manufacturing*, on-line and in press, no. DOI 10.1109/TSM.2018.2857818.

Zou H. and Hastie T., 2005. *Regularization and Variable Selection via the Elastic Net. Journal of the Royal Statistical Society, Series B*, 301–320.

REGRESSION ANALYSIS ON DATA RECEIVED FROM MODULAR IOT SYSTEM

Kristina I. Dineva
Tatiana V. Atanasova
Department of Modelling and Optimization
Institute of Information and Communication Technologies - the Bulgarian Academy of Sciences
1113, Sofia,
Bulgaria
E-mail: k.dineva@iit.bas.bg; atanasova@iit.bas.bg

KEYWORDS
IoT data, feature selection, regression, model, prediction.

ABSTRACT

This paper is aimed at choosing a model to determine in the short term the amount of honey in beehives based on historical data. The data is gathered from IoT-based monitoring system with modular architecture, which consists of sensors being located inside and outside the bee hives. The proposed solution enables better decision making by providing forecast results, thus beekeepers are able to plan their logistics to visit their apiaries. Regression analysis is used as techniques for modeling and analysis of variables in the forecasting model. Feature selection is provided as knowing the data is one of the first steps that must be taken before choosing the variables in the model. Predicting results are realized by using the model.

INTRODUCTION

Increasingly, recent technologies pay attention to the Internet of Things (IoT) as a tool that contributes to sustainable development and environmental protection. The latest innovative applications of the IoT help to improve the bee growing process and thus increase the quality, quantity, sustainability and economic efficiency of beekeeping. This paper aims at choosing a model to determine in the short term the amount of honey in beehives based on IoT historical data. The system with a modular architecture proposed in (Dineva and Atanasova 2017) combines all the benefits of current monitoring systems and adds additional functionality which helps to collect real-time and useful data by exploiting of smart sensors. Regression analysis is proposed to construct a model for predicting important values - indicators in the hive, which could be used to plan beekeepers' activities.

REGRESSION ANALYSIS

Regression analysis is a set of statistical methods for assessing the nature of relationships between variables. It includes many techniques for modeling and analyzing variables when the focus is on the relationship between a dependent variable and independent variables (predictors). Regression analysis means how the value of the dependent variable changes when any of the independent variables change while the other independent variables are fixed. Feature selection is provided as knowing the data is one of the first steps that must be taken before choosing the variables in the model.

Feature selection

Feature selection is one of the main concepts in machine learning, which has a significant impact on the model's performance.
It is necessary to remove the insignificant or partly relevant features because they may have a negative impact on the efficiency of the model. The presence of inappropriate data features reduces the accuracy of the model, changes the desired output, and is trained on inappropriate features. Choosing the right features before data modeling helps:
- Reducing excessive load - Less noise;
- Improving accuracy
- Reduce learning time - fewer data points lead to a reduction in the complexity of algorithms and faster training

Pearson Correlation
Pearson's correlation coefficient is a statistical model known as *r* value. For each two variables, a value is displayed that shows the correlation strength.
The coefficient is not affected by the magnitude changes of the two variables.

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \qquad (1)$$

where n - number of pairs of scores, $\sum xy$ - sum of the products of paired scores, $\sum x$ - sum of x scores, $\sum y$ - sum of y scores, $\sum x^2$ - sum of squared x scores, $\sum y^2$ - sum of squared y scores.
The two variables are quantitative and continuous.
Spearman Correlation
Spearman's coefficient expresses the extent to which two variables are monotonically linked. Spearman rank correlation r_s is designed for use on nonparametric and unbalanced data.

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} \qquad (2)$$

r_s – Spearman Rank correlation coefficient,
$\sum d^2$ - sum of the squared differences between ranks,
n – number of pairs of observations in the sample.
Kendall Correlation

Kendall's correlation measures the relationship between the rankings of different ordinal variables or different rankings of the same variable.
The use of rank correlation procedures provides the researcher with the opportunity to determine the ratios of not only quantitative, but also qualitative characteristics, in the case, if the latter can be ranked. Of course, the results are less accurate, since not the quantitative characteristics themselves are to be compared, but only the orders of their succession to each other.
Kendall's correlation is a nonparametric test that measures the power of dependence between two variables. If we look at two samples, *a* and *b*, where each sample size is *n*, we know that the total number of pairs with *b* is *n (n-1)/2*.

$$\tau = \frac{n_c - n_d}{\frac{1}{2}n(n-1)} \qquad (3)$$

n_c - number of concordant,
n_d - number of discordant.
Chi Squared
The bi-directional chi-squared test is a statistical method that measures how close the expected values are to the actual results. The method assumes that the variables are random and are derived from an adequate sample of independent variables. The resulting chi squared statistics shows how far the results of the expected (random) result are.

$$\chi^2 = \sum \frac{(O-E)^2}{E} \qquad (4)$$

O - the frequencies observed
E – the frequencies expected
Fisher Score
The Fisher method is an information estimate representing the amount of information a variable provides for an unknown parameter.
The result is calculated by measuring the difference between the expected value of the information and the observed value.

Application of feature selection methods

Knowledge of the data is one of the first steps to be taken (Tashev and Hristov, 2000) before choosing the variables to be included in the process. Given this, in the "Filter-based filter selection" stage, knowledge of the data plays an important role in the choice of method for modeling.
With the highest coefficient of influence, it is important to be the variable that actually has the greatest impact on the final result. And the coefficients with the result of 0 and ≈0 (near 0) are the variables that affect the least.
Through these factors, learning is more complete and higher accuracy is achieved because the coefficient of each variable of the selected method is taken into account. In order to achieve high accuracy, noise reduction and learning time, we set the desired number of functions to be 5 out of a total of 9 contained in a dataset.
After applying the various methods (Table 1) for features selection, it was found that the most appropriate method for the purposes of the study was Pearson's correlation, which best describes the degree of influence of the variables. It shows that the variables are quantitative, have a normal distribution, and the relationship between them is linear.

Table 1: Features Selection Results

	Tempe-rature	**hPa**	**Humi-dity**	**Wind Speed**	**Hour**
Pearson	0,58	0,44	0,28	0,09	0,01
Kendall	0,33	0,45	0,18	0,029	0,001
Chi Squared	146,86	235,20	97,07	40,84	2,29
Fisher	2,45	1,06	1,84	0,92	0,91
Spearman	0,47	0,61	0,26	0,038	0,002

The results obtained (Figure 1) show that the variables temperature, atmospheric pressure, humidity, wind speed and time are the most determinant factors, with a different influence on the change in the amount of honey in the beehive.

Figure 1: The Results Obtained

WORKING PROCESS

After the selection of columns for the work from the database, the normalization of the data, the clearing of the missing data and the selection of features for the work, four models are trained.
For training and testing of selected models, cross-validation is applied. With it, we create a *k* number of divisions of data and, respectively, a *k* number of folds. Cross-validation uses stratification by which we take data from different places of the dataset, thereby ensuring that each fold is a good representative of all information.
Evaluate model is a final step in the entire workflow that calculates a set of standard metrics for evaluation that show the error rate for regression models. Microsfot Azure Machine Learning Studio has being used (https://azure.microsoft.com) for the research (Figure 2).

INDICATORS FOR EVALUATION ON THE REGRESSION MODELS

The most important benchmarks for the model are: mean absolute error, root-mean-square error, relative absolute error, relative square error, and determination coefficient. Depending on the values obtained, these indicators are a base for selection the best model to represent the data. The model is considered to fit the data well if the difference between the observed and predicted values is small. The mean absolute error (MAE) measures how close the forecasts are to the actual results - so a lower score is better:

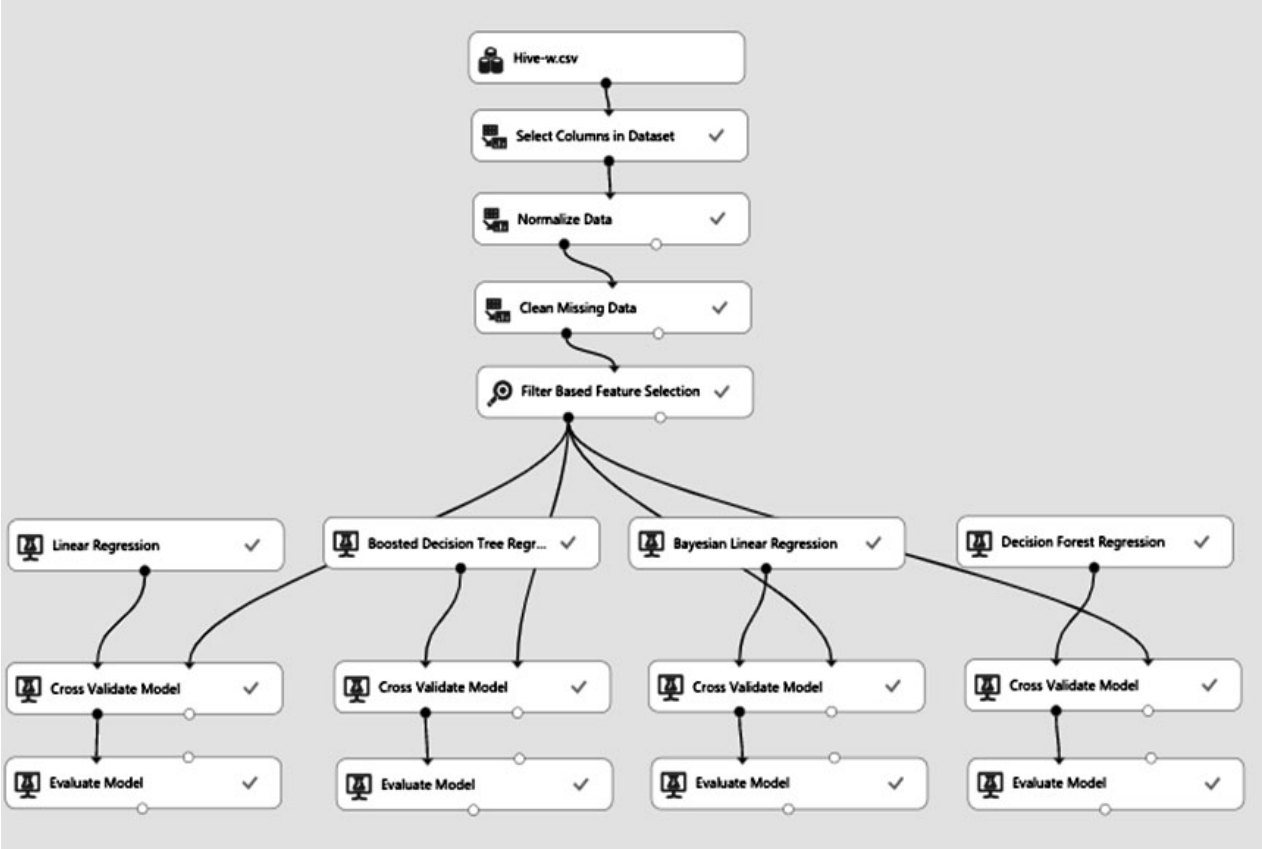


Figure 2: Workflow

$$MAE = \frac{\sum_{i=1}^{n}|p_i - a_i|}{n} \qquad (5)$$

The Root Mean Square Error (RMSE) creates a single value that summarizes the error in the model. Through equalization the difference, the metric ignores the distinction between over-prediction and inadequate prediction:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(p_i - a_i)^2}{n}} \qquad (6)$$

a - actual target,
p - predicted target.

The relative absolute error (RAE) is the relative absolute difference between the expected and the actual values. It is relative because the average difference is divided by the arithmetic mean:

$$RAE = \frac{\sum_{i=1}^{n}|p_i - a_i|}{\sum_{i=1}^{n}|\bar{a} - a_i|} \qquad (7)$$

The relative squared error (RSE) similarly normalizes the total squared error of the predicted values by dividing the total squared error by the true values.

$$RSE = \frac{\sum_{i=1}^{n}(p_i - a_i)^2}{\sum_{i=1}^{n}(\bar{a} - a_i)^2} \qquad (8)$$

The coefficient of determination (R^2) represents the predictive force of the model as a value between 0 and 1. Zero means that the model is random (explains nothing); 1 means that it has a perfect fit. However, care must be taken when interpreting R^2 values as low values can be completely normal and high values may be suspicious.

If the regression model is "perfect", SSE is zero, and R^2 is 1. If the regression model is a total failure, SSE is equal to SST, no variance is explained by regression, and R^2 is zero.

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \qquad (9)$$

$$SST = \sum(y - \bar{y})^2 \qquad (10)$$

$$SSR = \sum(y' - \bar{y}')^2 \qquad (11)$$

$$SSE = \sum(y - y')^2 \qquad (12)$$

R^2 – coefficient of determination
SST- Sum of Squares Total
SSR – Sum of Squares Regression
SSE - Sum of Squares Error

RESULTS OF A COMPLETED PROCESS

The results obtained after training and testing of the selected models are the determining factor for the correct model selection. Table 2 shows the resulting metrics after training and testing four regression models

The tabularly presented comparison of the metrics obtained for each applied regression model shows that linear regression is the model that best describes the statistical relationship between the variables under

Table 2: Resulting Metrics after Training and Testing Four Regression Models

Metrics	Linear Regression	Boosted Decision Tree Regression	Bayesian Linear Regression	Decision Forest Regression
Mean Absolute Error	**0.031631**	0.100482	0.066806	0.087543
Root Mean Squared Error	**0.092386**	0.145703	0.098254	0.138833
Relative Absolute Error	**0.229573**	0.729287	0.484872	0.635377
Relative Squared Error	**0.209455**	0.520975	0.236910	0.473003
Coefficient of Determination	**0.790545**	0.479025	0.763090	0.526997

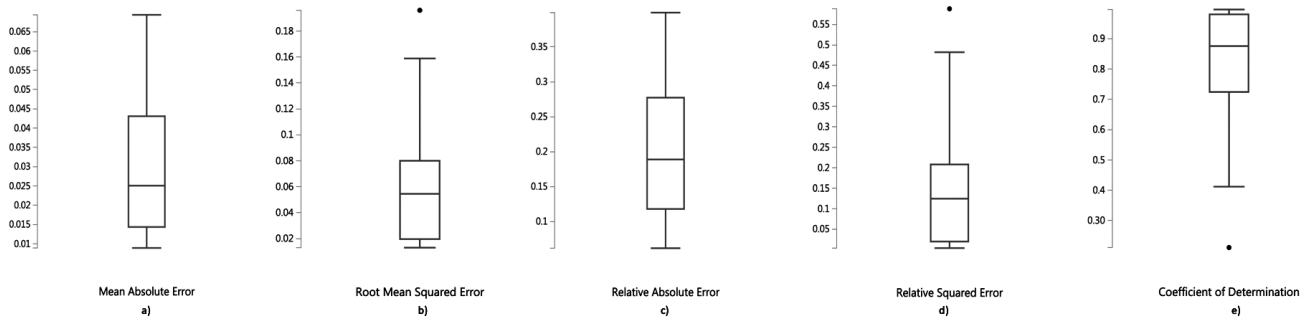


Figure 3: BoxPlot of Linear Regression Metrics

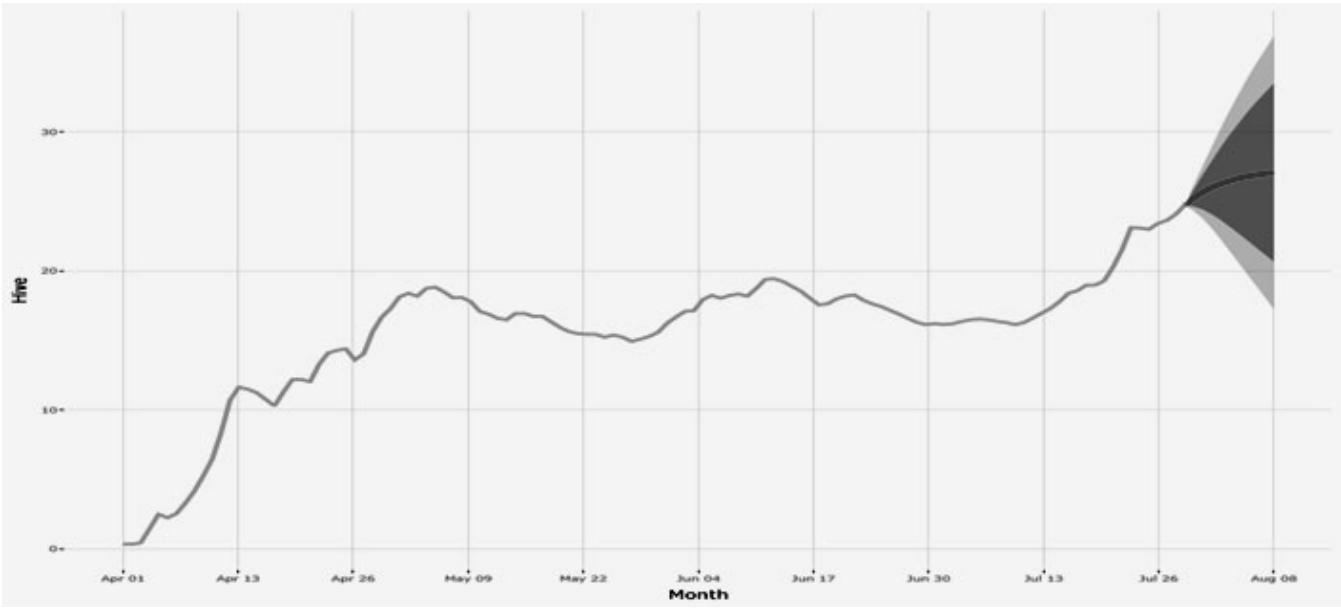


Figure 4: Short-term Forecast of Quantities

consideration. The absolute error (MAE) for linear regression prediction is 0.031631, which represents the mean value of the absolute difference between actual data points and the predicted result (Figure 3a).

The mean square error (RMSE) for linear regression is 0.0292386, which shows the default deviation of forecast errors. The value reflects the remoteness of the regression line data (Figure 3b).

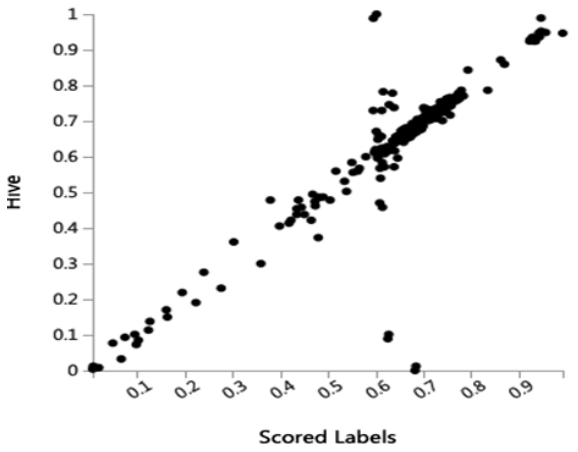


Figure 5: Dependence between *x* and *y*

The relative absolute error (RAE) for linear regression is 0.2229573, indicating the discrepancy between the exact value and the approximation (Figure 3c).
Relative squared error (RSE) for linear regression is 0.209455, indicating the mean value of the square error of the actual values (Figure 3d).
The coefficient of determination (R^2) for linear regression is 0.790545, which reflects the extent to which the dependent variable is predictable (Figure 3e). It can also be interpreted as the proportion of variance in the dependent variable predictable by the independent variable.
Figure 5 shows a linear relationship between x and y, which is maintained throughout the regression line. The graph is used to determine the normal error distribution. Standard residue values are also used. The regression line is straight, indicating that the residuals have a normal distribution. Simple tools like linear regression can actually be quite powerful (Christopher and Raghavan 2008).

PREDICTION OF RESULTS IN THE SHORT TERM

Because of the specificity of the case under consideration, it is important to obtain the most accurate data in the short term. Forecasts that reflect a long period of time would be inappropriate and unrealistic.
Figure 5 reflects estimates over a 10-day period that plays a key role in the short-term planning of activities.
The utility of analytical results is directly proportional to the decisions that can be taken to translate the results into a format that users can act upon (Balabanov et al. 2016; Blagoev 2018). Business interpretation and data visualization are invaluable means to better decision making.

CONCLUSION

Trends show an increase in the dependency between modern technologies and the successful development of beekeeping and other agriculture-related industries. New technologies enable beekeepers to make their industry as efficient as possible by reducing costs, increasing and generating new sources of revenue.
The short-term and long-term benefits of bee-keeping monitoring are related to the greater benefits of using IoT technologies - to enable better decision making by providing real-time data access by connected devices - bee hives (Dineva and Atanasova 2018).
Thanks to the forecast results, beekeepers will be able to plan their logistics to visit their apiaries, saving their time and money. On the other hand, a good organization will lead to a smaller number of hive bee hives that adversely affect honeybees due to a change in their favorable microclimate that will change as a result of the opening of the beehive.

REFERENCES

Balabanov, T., Zankinski, I., Barova, M. 2016. Strategy for Individuals Distribution by Incident Nodes Participation in Star Topology of Distributed Evolutionary Algorithms, Cybernetics and Information Technologies, Volume 16, No 1, Sofia, pp.80-88.

Blagoev, I., 2018. Using R Programming Language for Processing of Large Data Sets, Proc. Int. Conf. Big Data, Knowledge and Control Systems Engineering – BdKCSE'2018, pp. 91-98.

Christopher, D., Raghavan, P. 2008. Information Retrieval, Cambridge University Press, p. 175.

Dineva, K., Atanasova, T., 2017. Model of Modular IoT-based Bee-Keeping System. European Simulation and Modelling Conference ESM'2017, EUROSIS-ETI, 404-406.

Dineva, K., Atanasova, T. 2018. ICT-based Beekeeping using IoT and Machine Learning. Distributed Computer and Communication Networks, 21-st International Conference, DCCN 2018, 919, Springer, 132-143.

Stankovic, J.A. Research directions for the internet of things. IEEE Internet Things J. 2014, 1, pp. 3–9.

Tashev T.D., Hristov H.R. On One Approach for Modification and Expansion of the Information Interaction Models. J. "Problems of Technical Cybernetics and Robotics", Sofia, Academic Publishing House "Prof. Marin Drinov", No 49, 2000, pp.78-87.

WEB REFERENCES

https://azure.microsoft.com/

AUTHOR BIOGRAPHIES

KRISTINA DINEVA started her PhD study at the IICT - BAS in 2017. Kristina has over 10 years solid professional experience with bee-keeping acquired in the family farms located in different areas in Bulgaria. Kristina has some successfully completed software projects related to the web and IoT development with various software technologies and programming languages. Her current interests include innovative applications of the IoT as a technology that contributes to sustainable development.

TATIANA ATANASOVA is an Assoc. Prof. at the IICT-BAS. Her research interests lie at the cross-section of information society technologies, knowledge-based systems and complex control systems.

Web-based Platform for Collaborative Model-Driven Software Developement for IoT Devices

Imen Ben Ida
SERCOM laboratory,
Polytechnic School of Tunisia, Carthage university, Tunisia
imen.benida@gmail.com

Mortadha Dahmani
Corail Technology Research Center,
Tunis, Tunisia

Mohamed Ghazel
IFSTTAR,
COSYS/ESTAS,
Villeneuve d'Ascq, France

Abderrazek Jemai,
SERCOM laboratory,
Polytechnic School of Tunisia, INSAT, Carthage university, Tunis, Tunisia

Keywords—MDD, IoT, collaborative development, code generation

Abstract – The promising potential of the connected devices and sensors is playing an important role in improving the quality of services in different fields. However, developing software applications in this context raises several technology challenges as it relies on the participation of different stakeholders and it requires the consideration of several skills. In this paper, we focus on providing the support of high-level abstraction of software development. The proposed solution addresses the separation of concerns and ensures the respect of domain specifications from the first steps of software development for an IoT device. Namely, a web-based platform is proposed to provide a collaborative and model-based software development for IoT devices.

I. INTRODUCTION

The evolution of pervasive devices and communication protocols has made the Internet of Things (IoT) a reality that finds more and more applications in various domains. A fundamental component in IoT is the software with which things can be connected and exchange data. In the context of IoT, software can be found at different layers such as the middleware layer and the applications layer [1].

A software solution in the context of IoT is a combination of domain specification, communication technologies, interconnected apps, that would function together as one smart system to monitor, track and store information for specific services. In this context, several important challenges need to be addressed to enable the rapid development of IoT applications [2].

Firstly, IoT application development is a multi-disciplined process where knowledges from multiple concerns intersect. The individuals involved in the application development should have varied set of skills. Some of the involved disciplines are product definition and specifications, software, electronics, network specifications, hardware, security, mechanical integration and others. There is a need for easier and faster multi-discipline coordination between all the development teams. Secondly, IoT applications are executed on a network consisting of different devices in terms of types, interaction modes (e.g. Publish/Subscribe, Request/Response, Command), user platforms and user needs such as the signification and the priorities of collected data. Developers have to keep in mind these differences and the evolution issues both in the development phase and deployment phase.

Thirdly, the development of an IoT application for some fields such as healthcare, should take into consideration the domain specifications from the first steps of the development processes. The domain requirements are the guidelines to choose not only the data to be recorded or the type of sensors, but also the logic of the treatment process.

As a response to these challenges, recent researches have proposed the application of Model-Driven Development (MDD) approaches to IoT development following a variety of strategies to specify the system model, in which the heterogeneous elements are precisely identified. The use of models promotes the reuse of system elements across teams and other applications.[3] Added to that, it supports a rapid development of heterogenous applications with a minimal effort driven by the various stakeholders involved in the process.

In this paper, we propose a web-based platform that enables the participation of various stakeholders to introduce different types of models and to generate a code for a specific device in a personalized use case.

The rest of the paper is organized as follows: Section II presents the related works. The different components of the proposed platform are presented in Section III. In section IV, a use case in the healthcare field is illustrated. Finally, Section IV draws some concluding remarks and perspectives.

II. RELATED WORKS:

Recent works have proposed several approaches to improve the development process of IoT software. In [4], the authors aims to facilitate the learning of IoT development for novice developers by enabling them to easily become familiar with source code, written by other developers that faced similar issues. They propose an approach based on Code Recipes as documentation modules structured around code fragments, and which can be incorporated in various kind of tools such as Integrated Development Environment (IDE) extension.

The main contribution of [2] is a framework that separates IoT application development into different concerns based on a model driven development methodology and it integrates a set of high-level languages to specify them. The platform integrates code generation, task-mapping and linking techniques to provide automation. The evaluation based on two realistic scenarios shows that the solution generates a significant percentage of the total application code, which reduces development effort for IoT applications.

The authors of [6] propose a model driven methodology for the development of Web and mobile applications. The methodology is based on mapping the high-level description models using UML profiles into executable code and is

tested accordingly. It has been illustrated throw the creation of an application for visualizing energy data of residential homes. The obtained results prove that the methodology speeds up the application development process.
In [7] a prototype IDE for novices called CodePilot is presented. The tool enables multiple users to work together using a web-based programming session. In a unified system, they integrate real-time collaborative coding, testing, bug reporting, and version control management. The proposed solution aims to unify the collaborative development workflow into a single IDE.
All the previous works aim to facilitate the development of IoT applications using collaborative approaches. However, they are for general purpose and do not take in consideration the domain specifications a reference model. There is a need of domain model integration from the first step of IoT software development in order to respect the use cases requirements and offering personalized solutions.

III. PROPOSED WEB-BASED PLATEFORM

Our work aims to raise the level of abstraction and to enable different participants collaboration for IoT software development. Raising the level of abstraction improves the communication between team members through model representations. Added to that reusing existing models increases the productivity and efficiency in the design and development, especially of complex systems.

The proposed web-based platform provides the software developer with the necessary code that respects the specifications introduced by the participants and the use case configurations. The different components of the web application are described below:

1) The involved participents:

The development of an appropriate software for a specific connected device requires the integration of different concerns.
We define four main concerns, namely domain specifications, hardware specifications, network specifications and software specifications.
Each concern is defined by a participant according to his skills. The clear identification of expectations and specialized skills of each type of participant, helps them to play their role effectively.

Table 1: The participants' roles

Participant	Roles
Domain specialist	The domain experts specify the vocabular and the rules for an application domain.
Software developer	The software specialists define the structure of an IoT application by specifying the software components and their relationships using the domain model. The developed components take in consideration the use case configuration to offer personalized solutions.
Hardware specialist	The hardware specialists offer templates of the basic configuration for the devices. This role can be the function of one of the software development team or can be a service obtained by consulting extern specialists.
Network specialist	The specialists develop different templates of communication protocols used to connect the devices.

2) Model-based architecture:

The proposed architecture presented in Figure 2 illustrates the different concerns that support the specifications of each participant. It also presents the main steps to generate a software program for a specific device. The use case configurations are used to ensure the personalization of the generated code.

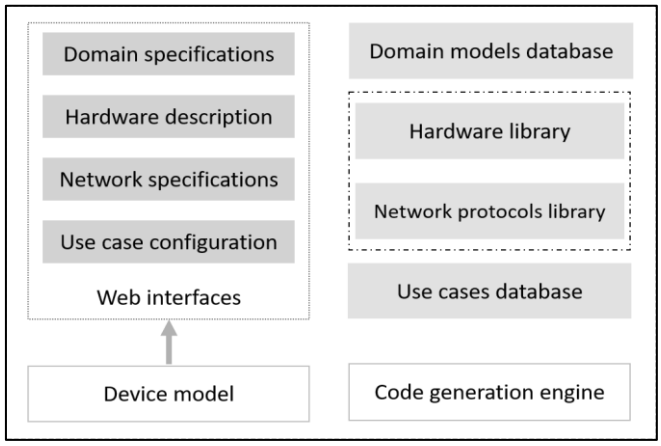


Figure 1: the model-driven approach

a) Device model

To ensure the interoperability of the eventual proposed domain models, we define a model presented in Figure 2. It presents the main proprieties of a device and its measurements in an IoT environment. It is considered as a reference model by all the developed interfaces described in the next paragraphs.

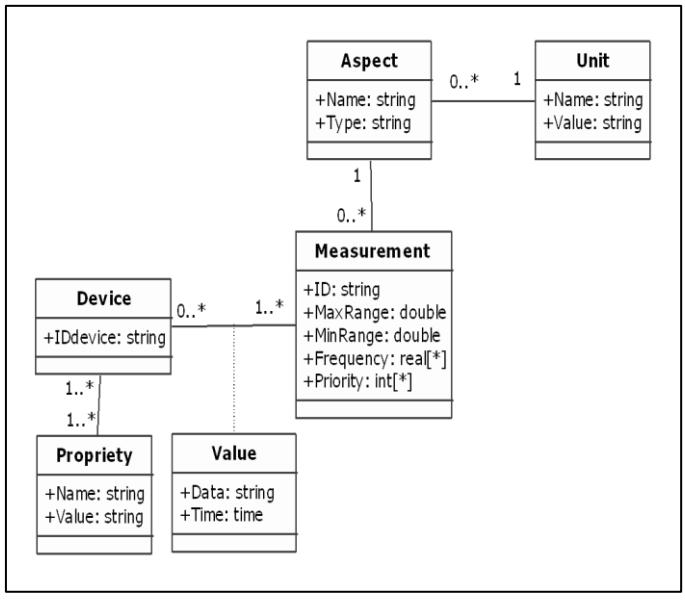


Figure 2: The device model

The main advantage of the device model is that it can be reused across applications of the same domain.
Our model presents the relationship between the following classes:

- Device: is an object that is able to interact electronically with its environment to accomplish a particular task. It is a producer and/or consumer of data that are exchanged in the network.
- Propriety: is a characteristic of the device. For example, the producer name or the vender addresses.
- Aspect: is a domain related data such as the temperature and the pressure. The class contains the name, the type and the unit of the data.
- Measurement: is a set of configurations for a specific aspect. It may support different priorities that depend on the case of application. The frequency attribute presents the possible intervals of collecting data. For example, a body temperature device may measure the data every hour or every minute depending on the patient situation. Each measurement must respect a range of values between the predefined MinRange and MaxRange attributes.
- Value : It is the sensed data at a specific time respecting a given measurement configuration.

b) Domain Specifications:

The first step of the proposed approach consists in the translation of the domain requirements to a domain model using the device model.
The proposed application gives the domain specialists the possibility of introducing specific values to instantiate a personalized device model. The Figure 3 presents the domain specialists interface. Each domain specification is saved as new measurement configuration in the domain models database. The corresponding Id of the new measurement will be added to the measurements list in the use case configuration interface.

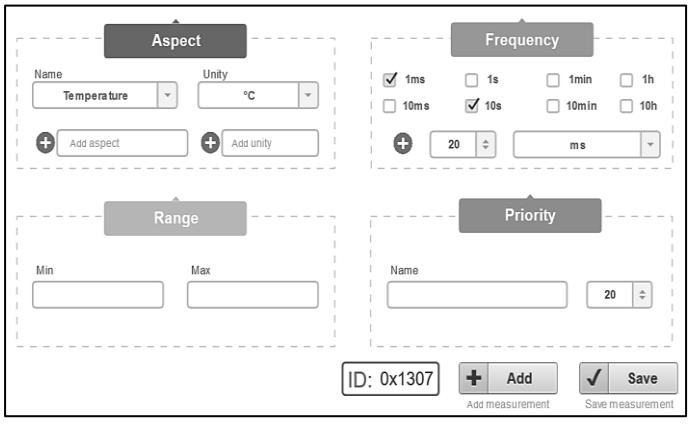


Figure 3: the domain specialists' interface.

c) Hardware specifications:

The platform allows the hardware specialists to introduce templates containing the code to be introduced into the devices depending on the hardware specifications. These templates are stored in the platform hardware library. The introduced code respects the device model.
For each new development of an IoT device software, the software developer uses the pre-defined templates depending on the target device.

d) Network specifications:

A Developer has to take into consideration not only the software and the hardware specifications of an IoT device, he should also consider the implementation of the network protocol used to exchange of information via the network. The same sensed data could be sent in different ways depending on the protocol specifications. The platform allows the network specialists to add the descriptions of network protocols and the necessary functions that must be used to send or receive data. These configurations are used to facilitate the software development of IoT devices and the update of the platform network library with new protocols descriptions.

e) Use case configurations

Using the domain model, the software developers modify the corresponding proprieties depending on the use case needs. First, he chooses the desired measurement and the he introduces the priority and the frequency of data collection.
Second, He select the network protocol. The platform proposes a list of devices from the hardware library. The, the developer chooses the device and validate the configuration.

a) Code-generation

After the use case configuration step, the code generation engine makes the link between the chosen network protocol template, the hardware specification and the specific proprieties related to the domain.

3) Use case in Healthcare domain

As an application domain, we chose to show the benefit of model-driven development of an IoT application in terms of personalization and reusability for patients with chronic diseases.
One of the main challenges in the management of chronic diseases is the efficient and continuous monitoring of the patient's health by the health professionals. [8]
On the other hand, continuous monitoring depends on the personalized situation of each patient which may considerably vary according to the type of disease and the patient context. For example, the frequency of collecting the heart beats changes depending on the patient illness. Taking these requirements in consideration from the first steps of developing a software for an IoT device makes the application more patient–centered and efficient.
The domain specifications are used to display the list of possible configurations such as the saved measurements identifiers. In our use case, the software developer chooses a temperature measurement using the interface illustrated in figure 4 and he defines the necessary configurations to generate the code.
The use case configuration ensures the personalization of the frequency of time used to collect temperature data regarding the patient condition.

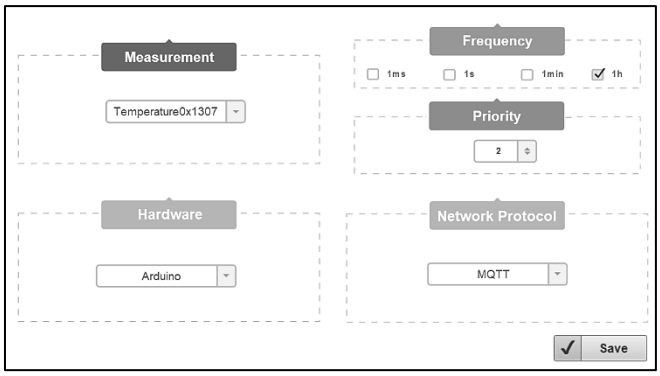


Figure 4: Use case configuration interface

For the hardware, we consider the case when an Arduino [9] device is used to be connected through internet and to send the temperature data with MQTT an network protocol [10]. MQTT protocol is based on a published subscribed pattern. The concerned aspect name introduced by the domain specialists is transformed into a topic by the code generator. The code presented in figure 5 shows the result of code generation after the validation of the use case configuration.

```
#include <ESP8266WiFi.h>
#include <PubSubClient.h>
// Update these with values suitable for your network.
const char *ssid =  "xxxxxxxx";
const char *pass =  "yyyyyyyy";
// These values are suitable for the use case configuration
const char *Aspect = aspect;
const char *Priority =  priority;
const char *DelayTime = Frequency;
const char *MaxVal =    MaxRange;
const char *MinVal =    MinRange;

// Network protocol
void callback(const MQTT::Publish& pub) {
  // handle message arrived
}
WiFiClient wclient;
PubSubClient client(wclient, server);
void setup() {
  // Setup console
  Serial.begin(115200);
  delay(DelayTime);
  Serial.println();
  client.set_callback(callback);
}
void loop() {
  if (WiFi.status() != WL_CONNECTED) {
      WiFi.begin(ssid, pass);
    if (WiFi.waitForConnectResult() != WL_CONNECTED)
      return;
    Serial.println("WiFi connected");
  }
  if (WiFi.status() == WL_CONNECTED) {
    if (!client.connected()) {
      if (client.connect("arduinoClient")) {
        if (data > MinVal && data < MaxVal){
          publish(Measurement,data);
        }
        else{
          publish(Measurement,"alert");
        }
    delay(DelayTime);
      }
    }
    if (client.connected())
      client.loop();
  }
}
```

Figure 5: The generated code for the healthcare use case

In the first part of the generated code, which is the declaration of the variables, is independent from the network protocol and it respects the device model.

In the declaration of MQTT client code, the network specialist uses the term measurement for the topic name.

IV .CONCLUSION AND PRESPECTIVES

In this paper, we focus on providing support for high-level abstraction of software development to address the separation of concerns and ensure the respect of domain specifications from the first steps of development. A model-based approach is proposed to describe the use of different specifications. As an application of the approach, a healthcare use case is presented.

Its main role is to reduce the time and the complexity of the development of software for IoT. We achieve this by separating IoT application development into different concerns and by integrating a set of web technologies to ensure the participation of the specialists of each concern.

Finally, we implement a part of the proposed solution as a web application. The proposed application gives the possibility to introduce the user configuration. As future work, we plan to add features that allow the different stakeholders to introduce their specifications. The implemented web application will support not only the user configurations as parameters but also the modification of other concerns.

REFERENCES

[1] S. Simon, F. B. A., and V. Thomas, “Designing an application store for the internet of things: Requirements and challenges,” in Ambient Intelligence: 12th European Conference, AmI 2015, Athens, Greece, November 11-13, 2015, Proceedings, B. De Ruyter, A. Kameas, P. Chatzimisios, and I. Mavrommati,Eds. Springer International Publishing, 2015, pp. 313–327.

[2] Pankesh Patel, Damien Cassou, Enabling High-level ApplicationDevelopment for the Internet of Things, The Journal of Systems & Software (2015), doi:10.1016/j.jss.2015.01.027

[3] R. France and B. Rumpe, "Model-driven Development of Complex Software: A Research Roadmap," Future of Software Engineering (FOSE '07), Minneapolis, MN, 2007, pp. 37-54. doi: 10.1109/FOSE.2007.14

[4] Corno, Fulvio; De Russis, Luigi; Sáenz, Juan Pablo (2018). Easing IoT Development for Novice Programmers Through Code Recipes. In: 40th International Conference on Software Engineering:Software Engineering Education and Training Track, Gothenburg, Sweden, May 27-June 3 2018

[5] Pal A., Mukherjee A., P. B. (2015) Model-Driven Development for Internet of Things: Towards Easing the Concerns of Application Developers. In: Giaffreda R. et al. (eds) Internet of Things. User-Centric IoT. IoT360 2014. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 150. Springer, Cham

[6] E. Ebeid, M. Valov and R. H. Jacobsen, "Model-Driven Design Approach for Building Smart Grid Applications," 2016 Euromicro Conference on Digital System Design(DSD), Limassol, 2016, pp. 260-267.doi: 10.1109/DSD.2016.94

[7] Jeremy Warner and Philip J. Guo. 2017. CodePilot: Scaffolding End-to-End Collaborative Software Development for Novice Programmers. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI ’17). ACM,New York, NY, USA, 1136–1141. https://doi.org/10.1145/3025453.3025876

[8] Cristiano André da Costa, Cristian F. Pasluosta, Björn Eskofier, Denise Bandeira da Silva, Rodrigo da Rosa Righi, Internet of Health Things: Toward intelligent vital signs monitoring in hospital wards, Artificial Intelligence in Medicine, Volume 89, 2018, Pages 61-69, ISSN 0933-3657, https://doi.org/10.1016/j.artmed.2018.05.005.

[9] https://www.arduino.cc/

[10] Banks, Andrew et GUPTA, Rahul. *MQTT Version 3.1. 1*. OASIS standard, 2014, vol. 2

AN OVERVIEW OF VIRTUAL AND AUGMENTED REALITIES IN STEM EDUCATION

Plamen D. Petrov and Tatiana V. Atanasova
Department of Modelling and Optimization
Institute of Information and Communication Technologies - Bulgarian Academy of Sciences
1113, Sofia,
Bulgaria
E-mail: p.petrov@iit.bas.bg; atanasova@iit.bas.bg

KEYWORDS
Virtual reality, Augmented reality, Hologram, STEM education.

ABSTRACT

Two of the most popular 3D visualization and modelling technologies with haptic and touch feedback possibilities are analysed. These include virtual and augmented reality tools, as well as a specific solution, incorporating augmented reality. A case study for delivering STEM (Science, Technology, Engineering and Mathematics) content using this tool at one secondary school in Sofia is presented. The experience gained for one school year of using facilities for a STEM enrichment program has been examined.

INTRODUCTION

We are living in a time of rapid advances in terms of both the capabilities and the cost of virtual reality, multi-user virtual environments, and various forms of mixed reality. These new media potentially offer extraordinary opportunities for enhancing both motivation and learning across a range of subject areas, student developmental levels, and educational settings. With the development of practical and affordable virtual reality and mixed reality, people now have the chance to experience immersive learning both in classrooms and informally in homes, libraries, and community centres (Liu et al 2017).
The technology, used in the classroom, must be adaptable, responsive, immersive and engaging, individualized and appropriate. It should provide the learner with unique experience that is not easily replicated or possible in the traditional classroom.
The technology, especially when used in STEM education, should allow students to explore and interact with STEM content and to practice, not worrying about economical or ethical issues like costly consumables or animal injury or killing. It gives opportunities for experiments and recovers from failure, whilst working in a safe environment. And virtual, augmented and mixed reality tools can do just that.

UPTODATE APPROACHES

Virtual reality (VR) is an artificial environment which is experienced through sensory stimuli (such as sights and sounds) provided by a computer and in which one's actions partially determine what happens in the environment (Merriam-Webster Dictionary). This is the use of computer technology to create a simulated environment.

Virtual Reality's most often used component is the head-mounted display (HMD). Audio-visual information is most easily replicated in Virtual Reality, but a lot of research and development efforts are being conducted into the other senses. Tactile inputs allow users to feel as though they're actually walking through a simulation, rather than sitting in a chair. Haptic technologies, also known as feedback tech, have progressed from simple spinning-weight motors to high quality ultrasound technology. It is now possible to hear and feel true-to-life sensations along with visual VR experiences.

A lot of educational technology companies are using virtual reality to bring true to life experiences to the classroom, while highlighting the technology's ability to inspire and grab the attention of the students. In the area of STEM education, fully interactive virtual lab simulations are designed to engage and stimulate a student's natural curiosity as they learn.

Labster, a Copenhagen-based company plans to use VR to bring the science lab to life for students. Labster's immersive education tools and virtual labs are used in about 150 universities worldwide. Now, it is developing a new VR product range in partnership with Google. The company aims to engage students through "storytelling, narrative and interactive gaming elements". VR offers a physical feeling of being in the lab, which gives students confidence when they progress to an actual lab environment (Kennedy 2018).

But there are certain risks, too. Experts are still trying to understand the impact of VR on children's learning. Research conducted by Bailenson and his team at their Stanford lab in 2008 looked at the potential psychological effects on young children using VR. In some cases, children who experienced swimming with whales in a VR environment developed false memories of having visited SeaWorld in real life (Oh and Bailenson 2017; Kennedy 2018).

Bailenson believes that concerns about VR use can be addressed in two ways: moderation and supervision. VR

is an engaging tool, but within a 40-minute lesson plan, it should be used carefully – two or three minutes.

Augmented reality (AR) is the integration of digital information with the user's environment in real time. Unlike virtual reality, which creates a totally artificial environment, augmented reality uses the existing environment and overlays new information on top of it (whatis.techtarget.com).

The popularity of AR is rising because it brings elements of the virtual world into the real world, thus enhancing the things we see, hear and feel. It brings computer generated objects into the real world that can be only seen by the user. When using augmented reality applications, the user sees combination of synthetic and natural light. Overlaying images are projected on top of a pair of glasses, which allows the images and interactive virtual objects to be on top of the user's view of the real world. Two of the ways AR can be experienced are through headsets that people wear and through displays on mobile devices. Budget-priced headsets that can be used with smartphones and tablets are available. On the other side are the high-end Microsoft's HoloLens and Facebook's Oculus VR headset. One of the things that can make AR accessible is the AR Applications. AR could be very suitable for simulations, especially in the area of STEM education.

With the rise in computing power and lowering of computer costs has been an increase in the use of simulations. A simulation, for the purposes of this study, is a computer-based interactive environment with an underlying model. In the STEM field in particular, real equipment can be difficult to obtain, so simulations enable students to experience phenomena they normally would not be able to experience first-hand. For example, simulations can take the place of laboratory equipment that might be too expensive or dangerous to have in a school. Simulations can also be used to explore phenomena that occur over long or extremely short time periods in a way that can easily fit into a class period. With simulations, students can also manipulate variables and see the results of multiple experiments without having to actually replicate them. In view of these benefits, it is believed that using simulations in the classroom can help improve learning (D'Angelo et al 2013).

CASE STUDY

One of the most successful examples of AR system is zSpace (zspace.com). In 2015, zSpace Inc. released the All-in-One solution for Education consisting of a virtual reality monitor and computer. The system provides students with a realistic learning environment that aligns to the Next Generation Science Standards (NGSS), Common Core and state standards. The virtual-holographic images can be lifted from the screen and manipulated with the accompanying stylus. The solution allows for group interaction as participants only need 3D glasses as opposed to VR headsets (Do 2015).

The zSpace system uses a 1920 x 1080 pixel LCD screen, and the hardware switches between the left and right images through a circularly polarized light that enters the eye. The glasses contain small reflective tabs that the computer uses to track where users are looking. This software prevents nausea and headaches by keeping the image focused.

zSpace combines elements of AR and VR to create lifelike experiences that are immersive and interactive. The system lets the students dissect animals and organs, manipulate compounds and structures in 3D and run experiments without the need for costly lab equipment and supplies.

zSpace was chosen as the AR tool for STEM education at the 134 Dimcho Debelianov School in Sofia, Bulgaria, in 2018 (https://www.ort.org/news/opening-new-stem-laboratories-gives-sofia-pupils-room-inspiration).

Figure 1: Facilities for a STEM Enrichment Program

The key factors when choosing the solution were:

- Scope – the technology can be used in many subjects and can address different problems;
- Collaboration – the possibility for work in small groups of 2-3 students; the technology itself encourages the teamwork;
- Experimentation – opportunities to perform scientific procedures;
- Creativity – students' creativity is encouraged and they can free their imagination and take risk without fear of failure;
- Critical thinking and evaluation – the technology supports complex problem solving and evaluative thinking, helping them to analyse and better understand abstract concepts.

Such type of technology was new for the school and one of the aims of the teachers was to measure the system's effectiveness. That's why the model that was used included three important elements (Figure 2):

- Well-defined learning objectives;
- Authentic learning experiences;

- Evaluation – students' results and their learning experience.

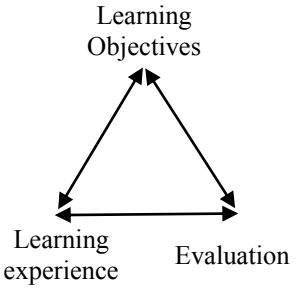


Figure 2: Elements to Measure the System's Effectiveness

Among zSpace learning and simulation applications are: zSpace Studio – a set of pre-made activities, available in different languages; Newton's Park – physics simulation environment, promoting discovery through experimentation and allowing data gathering and analysis; Human Anatomy Atlas – for independent student research on the human body, allowing at the same time self-assessment of learning.

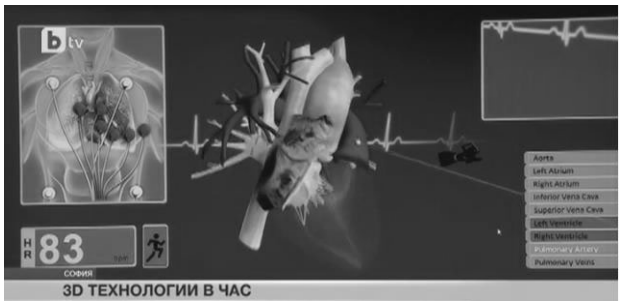


Figure 3: The Biology Laboratory

The experience gained for one school year has shown that the virtual experiments led to much more advanced learning than would ever be possible in a typical classroom. These initial findings have shown that simulations are a promising way for improving students' learning outcomes, especially in STEM subjects. It was noticed however that further qualitative and quantitative analysis is required.

Evaluating whether learning objectives are being met, especially when a lot of investments in equipment have been done, is crucial for the effective usage of this equipment. This requires creating an evaluation strategy, collecting assessment data and making recommendations for improvements. A special tool, A-Class HiTeach, is used for collecting data and validation of the current training strategy and students' achievements. It uses cloud technologies and artificial intelligence to generate data-driven recommendations for improvements.

CONCLUSION

Experience has shown that whatever applications and tools are available; it is difficult for people to become more educated. It is clear that human and technology must go together, but the "golden section" between the human and technology must be sought. Hologram is one such possible solution. In the process of exploring the human-technology relationship, the science fiction can be very useful. Many books have remained only a fiction but some of them have become true scientific insights.

ACKNOWLEDGEMENTS

Research described in this article was partially supported by the National Scientific Program "Information and Communication Technologies for a Single Digital Market in Science, Education and Security (ICTinSES)", financed by the Ministry of Education and Science, approved by DCM # 577/17.08.2018.

REFERENCES

D'Angelo, C., Rutstein, D., Harris, C., Bernard, R., Borokhovski, E., Haertel, G. 2013. Simulations for STEM Learning: Systematic Review and Meta-Analysis (Executive Summary). Menlo Park, CA: SRI International.

Do, K. 2015. 3d Technology at ISTE 2015. *Cyber Science 3D.*

Kennedy E., 2018. Can virtual reality revolutionize education? CNN (November 1)

Liu, D., Dede, C., Huang, R., Richards, J. (Eds.) 2017.*Virtual, Augmented, and Mixed Realities in Education*, Springer Singapore.

Oh S. Y., J. Bailenson, 2017. Virtual and augmented reality, *The international encyclopedia of media effects*, John Wiley & Sons, Inc., 1-16.

WEB REFERENCES

https://edition.cnn.com/2018/11/01/health/virtual-reality-education/index.html

https://www.ort.org/news/opening-new-stem-laboratories-gives-sofia-pupils-room-inspiration/?fbclid=IwAR179SeAc_XuRQyJNXU_LddsROMHv-zmOYBXw6PLBI2ooRm-UKVNhTfVSaU

https://whatis.techtarget.com/definition/

https://www.merriam-webster.com/dictionary/

https://zspace.com

AUTHOR BIOGRAPHIES

PLAMEN PETROV obtained Master Degree in Mathematics and Informatics from Sofia University, "St. Kliment Ohridski". He started his PhD study at the IICT-BAS in 2019. Plamen has solid professional experience in school and university level teaching. He has co-authored several students' books in the area of ICT and has specialized in the field of educational technologies. His current interests include Instructional design, e-Learning Methods and Models and Learning Content Management Systems Development.

TATIANA ATANASOVA is an Assoc. Prof. in the Modelling and Optimization Department at the IICT – BAS. Her research interests lie at the cross-section of information society technologies, knowledge-based systems and complex control systems.

FINANCIAL DATA SIMULATION

INFORMATION TECHNOLOGY FOR STRUCTURAL FLOWSHEET SYNTHESIS, BASED ON THE CONSTRAINT SATISFACTION METHODS

Alexander Zuenko
Andrey Oleynik
Yurii Oleynik
FIC KSC RAS
Institute for Informatics and Mathematical Modelling
Apatity, Russia
E-mail: zuenko@iimm.ru

Valeriy Birukov
Roman Nikitin
FIC KSC RAS
Mining Institute
Apatity, Russia
E-mail: birukov@goi.kolasc.net.ru

KEYWORDS
Computer Integrated Manufacturing and Engineering, Computer performance, Industrial processes, Decision support systems, Optimization

ABSTRACT

The paper presents the initial phase of studies on the development of an information technology for structural synthesis of multistage flowsheets. The study is supposed to define and solve a structural synthesis problem in form of a constraint satisfaction problem. The methods and heuristics developed in the constraint satisfaction theory allow both quantitative and qualitative constraints in the subject domain to be jointly processed. Within the structural synthesis technology developed it is possible to consider both a great number of heterogeneous objective factors effecting the value of the target function of the object synthesized, and the subjective constraints. A "test object" to be synthesized, is a mineral processing flowsheet.
The study utilizes the principal of block (stage-by-stage) generation of flowsheet models. The use of stages as elements of the flowsheet decomposition allows simultaneous application of both technological and economical criteria of its effectiveness, as well as the different types of the system decomposition.

INTRODUCTION

The aim of the research presented in the paper is to develop an information technology for structural synthesis of effective schemes of multistadial technological processes, which are characterized by the presence of heterogeneous constraints on compatibility of the equipment implementing the process and the performance indicators. Most known approaches to the solution of the structural synthesis problem have no means for formal definition of heterogeneous constraints on combinations of structural elements (Alpatov 2019). So the task to develop a structural synthesis technology that will provide a formalized definition, an automated analysis and accounting constraints, is very actual.
The novelty of the developments within the project is in that the problem of structural synthesis is suggested to be defined and solved as a constraints satisfaction problem (CSP). In the presence of a complex of heterogeneous performance criteria and constraints, reducing the problem of structural synthesis of the multistadial technological processes schemes to the CSP, to the Constraint Optimization Problem (COP), to be exact, allows us to take advantage of an extensive set of methods and heuristics developed in the constraints satisfaction theory. The constraints programming systems architecture provides the prerequisites for joint processing both quantitative and qualitative constraints of the subject domain.
Chosen as a 'test object' is a mineral processing flowsheet. From the application point of view, the technology developed provides operational synthesis of effective structural schemes of production processes, with the composition and values of the performance criteria varied. It will contribute the solution of the engineering problem connected with the selection of the flowsheets topologies for modernized and newly created factories.
The functioning of a number of software packages for flowsheet modeling (JK Sim Met (Morrison and Richardson 2002), CHEMCAD, USIM PAC, a software package of Andritz Automation Inc for crushing-grinding – classification modeling (Brochot 2002), MODSIM (Herbst 2004)) has been analyzed. The analysis results show that the common disadvantage is inability to have automated generation of economically effective flowsheet topology in mineral processing.

THE CONSTRAINTS PROGRAMMING TECHNOLOGY AND ITS OPPORTUNITIES

The constraint programming technology allows us to represent the model of the problem in a declarative view in a language similar to the language of mathematics. To apply the technology considered, it is necessary to define an applied problem as a constraint satisfaction problem.

According to (Russel and Norvig 2010, Rossi et al. 2006), a constraint satisfaction problem consists of three components: $< X, D, C >$. X is a set of variables $\{X_1, X_2, ..., X_n\}$. D a set of domains$\{D_1, D_2, ..., D_n\}$, where D_i is the domain of variable X_i. C a set of constraints$\{C_1, C_2, ..., C_m\}$, which assign admissible combinations of the variables values.
Each domain D_i defines a set of admissible values$\{v_1, ..., v_k\}$ for variable X_i. Every constraint is a couple $< scope, rel >$, where *scope* is a set of variables which take part in constraint, *rel* is a relation regulating the admissible combinations of the values which can be taken by the variables from scope.
The notion of constraints is interpreted rather widely. Any relationship between the variables of a domain may be taken as a constraint. Any arithmetic expressions (linear equations/inequalities, in particular), logic formulae, tables, expressions formulated in the language of specialized theories (e.g. the graph theory, the finite automata theory, and the like) may be taken as constraints.
The solution of the CSP is a complete assignment satisfying all the constraints. In some cases, it is necessary to reach all the solutions. In other cases, it is necessary to reach such a solution in which the values of variables would optimize a specified functional. The CSPs are classified as NP-complete problems.
The constraints programming technology provides powerful and flexible methods, algorithms of solution of the combinatory search problems. The peculiarities of the technology are:

1. In the viewpoint of the end user, the CSP is formulated as a declarative problem, in the language similar to the language of mathematics. The order in which the constraints are set is not essential.

2. Any algorithm of constraints satisfaction should contain two obligatory components: a) a component realizing inference (propagation); b) a component realizing search.

3. Inference (propagation) is realized as a purposeful reduction of initially specified variable domain.

4. Heuristics used in search procedures, are universal, being not developed for a particular problem. The search procedures are, first of all, those of backtracking search.

5. Owing to the constraints programming systems architecture, there is a possibility to jointly process qualitative and quantitative constraint.

6. An opportunity is provided to support models open for operational modifications. In adding/removing constraints to/from a model, it not necessary to write new methods to solve a problem.

Constraints programming is especially useful when there is no possibility to reach an adequate solution with the tools of "usual" mathematics. It is used in solution of planning, design and prediction, in engineering and economical calculations, in graphical interface creation, in natural language understanding systems, etc.(Rossi et al. 2006, Margaux et al. 2017, Kreter et al. 2016, Buscemi and Montanari 2008) Among the most known international systems realizing the constraints programming paradigm we can point out Prolog III, CLP(R), CLP(BNR), clp(FD), CHIP, ILOG Solver, Newton, etc. In this study, the authors have used ECLiPSe Constraint Programming System (Apt and Wallace 2006). Beside the base arithmetic operations, the solver ECLiPSe allows us to use trigonometric, logarithmic and some other functions in defning CSPs. As you see below, the functions like $exp(x)$ are useful in modeling constraints for the problems considered in the paper. The ECLiPSe allows us to solve the constrained optimization problem by specifying certain criteria. In addition, there is a large set of global constraints, heuristics for the choice of variables and values of variables, which allow accelerating the execution of developed algorithms.

DEFINITION OF THE MINERAL PROCESSING FLOWSHEET SYNTHESIS PROBLEM AS A CONSTRAINTS SATISFACTION PROBLEM

The study is based on the principal of block (stage-by-stage) generation of mineral processing flowsheet models. The use of stages as elements of the technological process decomposition allows us to simultaneously apply both technological and economical criteria of its performance, to use, if necessary, a functional (by function), component (by the type of elements) and structural (by the type of relations between the elements) decomposition of the system.
The solution of the problem connected with making a structural synthesis within the constraints programming technology is reduced to the solution of a series of constraints satisfaction problems. Considered first are the constraints peculiar for the first block (stage) – the constraints satisfaction problem is defined. In case it is impossible to reach any solution, a new CSP is generated: added to the flowsheet is one more block (stage), formed also are constraints characterizing a newly generated flowsheet. As a rule, a maximum and minimum possible amount of blocks in the flowsheet may be assessed a priori by experts.
Defined below in detail are mineral processing and the effective flowsheet synthesis problem as a constraints satisfaction problem.
Based on the amount of the useful component, the mineral processing products may be classified into 3 groups:

- **tailings** is a product with minimum content of the

useful component, its further processing is useless;

- **concentrate** is a product with the useful component content adequate for further realization, i.e., the final product of mineral processing;
- **middlings** is a product with the intermediate useful component content which, as a rule, is to be further processed in order to be classified into the previous groups.

Processing is a combination of methods of metals or minerals separation, based on their physical and chemical properties. In the case studied, this process is a set of stages of mass flow processing (Fig. 1).

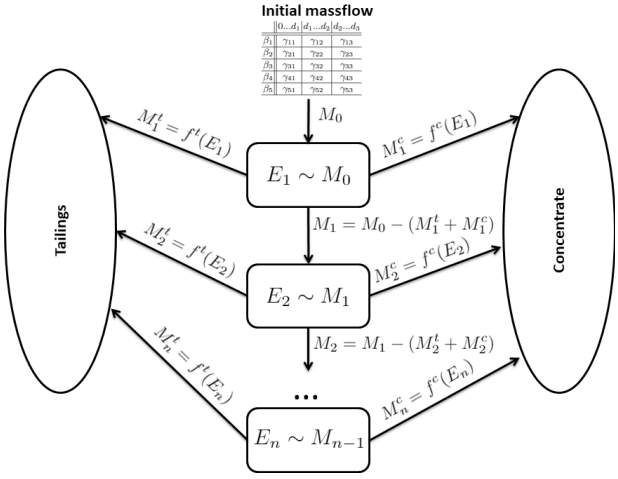


Figure 1: The Stages of Mass Flow Processing

The input mass flow (feed) M_0 is divided into 3 flows:
M_1^t is a part of the mass flow, being the tailings;
M_1^c is a part of the mass flow, being the concentrate proper;
M_1 is a part of the mass flow, being middlings that follows to the next stages of processing.
The separation of the input mass flow (feed) into tailings, concentrate and middlings depends on energy (E_1) spent in grinding. In turn, this energy depends on the characteristics of the input ore (strength and size of ore, in particular), mass flow volume and on the grain size fineness planned to be produced in the output of the structural block (stage).
Then, mass flow M_1 goes stage 2 where it is separated into 3 mass flows, and so on, up to the final stage where mass flow M_{n-1} is separated only into tailings and concentrate. Thus, we may define the following constraints:

$$M_0 = M_1 + M_1^t + M_1^c;$$
$$M_k = M_{k+1} + M_{k+1}^t + M_{k+1}^c, \quad k = 1, ..., n-2;$$
$$M_{n-1} = M_n^t + M_n^c;$$
$$M^t = \sum_{i=1}^{n} M_i^t; \quad M^c = \sum_{i=1}^{n} M_i^c; \quad M_0 = M^c + M^t;$$

In theory, based on the ore characteristics, it is possible to calculate such a grain size fineness for the mass flow to be divided into tailings and concentrate immediately, in one stage. It will demand huge amount of energy however. So, the process is carried out in several stages, with the amount of the mass flow being reduced at each next stage of processing. Also, the grain size fineness changing from stage to stage, depends on the classifier mesh dimension. The variants of this parameter values are known and finite, so these can be used as a domain of the corresponding parameter of the CSP. Thus, there is a finite set of mass flow grinding variants. Using the grinding variant selected it is possible to calculate the amount of energy necessary for the process, using the equation: $\frac{d_1}{d_2} = exp(k * \frac{E}{V})$, where d_1 is the ore diameter before grinding, d_2 – after grinding, V is the amount of the mass flow ground, E is the amount of energy spent for grinding, k is a constant coefficient depending on the ore properties. Taking the above said into account, it is possible to derive the condition of optimization, as well as the boundary value of the energy spent for all the process:

$$\sum E_i \to min, \quad E_i < E_o, \quad \sum E_i < E_o,$$

where E_o is the energy necessary for separation of whole the mass flow into tailings and concentrate immediately in one stage.
As a whole, each constraint satisfaction problem (the constrained optimization problem) from a set of CSPs which has been mentioned above, includes determining the sizes to which the mass flow should be ground at each stage in such a way that the total energy spent in all the flowsheet could be minimum. In addition, it is necessary to observe additional constraints, which dictate the amount of useful component that is allowed to be in tailings and in concentrate. The amount of possible stages in a flowsheet is limited (this study is supposed to have five stages at least).So, in this case, a series of constraint satisfaction problems consists of five problems. The results of the solution of each of the previous problems are used in defining the variables domains for each next problem.
In defining previous constraint satisfaction problems, we consider each stage of the flowsheet to be a so-called "black box", without making it clear in what way the characteristics of input and output mass flows are presented. In fact, each stage consists of the following processes: grinding, classification and separation. Given below is the defining of these processes. It should be noted here that each process changes the characteristics of the input mass flow. The characteristics of the mass flow are given in the Table 1.
The columns in the Table 1 characterize ore fractions are different in size. The quantity of the size fractions is determined/specified depending on the modeling accuracy desired. The rows of the Table 1 define the distribution

Table 1: The Characteristics of the Mass Flow

	$0...d_1$	$d_1...d_2$	$d_2...d_3$
β_1	γ_{11}	γ_{12}	γ_{13}
β_2	γ_{21}	γ_{22}	γ_{23}
β_3	γ_{31}	γ_{32}	γ_{33}
β_4	γ_{41}	γ_{42}	γ_{43}
β_5	γ_{51}	γ_{52}	γ_{53}

of the useful component content inside each fraction. The useful component content varies depending on the useful component content in waste to that which is possible at its maximum in the mass flow. In the Table 1 cells there is some number γ_{ij} showing what fraction of the total mass flow is made up by particles of the corresponding size, with the specified content of the useful component.

Grinding

To model grinding, masses of each ore fraction are calculated by the initial data from the Table 1. Then, the particles of each fraction are affected by the energy proportional to the mass of corresponding ore fraction.
Calculated for each fraction is a new table of mass fraction distributions. It is supposed that the result of each fraction grinding does not depend on the results of other fractions grinding. The tables of each fraction distribution are generated in accordance with the equation of "reflected normal distribution":

$$f(d) = \frac{1}{a\sqrt{2\pi}}\Big[exp\Big(-\frac{(d-d_2)^2}{2a^2}\Big) + exp\Big(-\frac{(d+d_2)^2}{2a^2}\Big)\Big],$$

where $a = 0.4 * (exp(-0.5E_{ism}) - (exp(-E_{ism})))$, $d_2 = \frac{d_j}{exp(E_{ism})}$, $E_{ism} = E_1 - m_{frac}$, m_{frac} is mass coefficient of the corresponding ore fraction, d_j is the size of the j-th fraction.
The distribution tables generated for each fraction are reduced again into a general table of the mass flow ground by energy E_1. Then, applied to this table is a model of mineral liberation in grinding.
To model the mineral particles liberation in grinding i.e., the formation of particles densities distribution in newly formed size classes, use is made of the formula of beta distribution density. Shown below is the example of the beta distribution density formula for 2-component ore, where β^+ is the useful component content.

$$f_{beta}(\beta^+) = \frac{1}{B(u,v)}(\beta^+)^{u-1}(1-\beta^+)^{v-1},$$

where $B(u,v) = \int_0^i t^{u-i}(1-t)^{u-i}dt$ are beta-functions or Euler integrals of the first kind; u and v are the parameters of the form of beta distribution.

$$u = \beta^+ * \Big(2\frac{d_j}{d_v}\Big), v = (1-\beta^+) * \Big(2\frac{d_j}{d_v}\Big).$$

The useful component is included into a mineral as isolated particles. The grain size of the useful component in ore determines "grain impregnation". Parameter $\frac{d_j}{d_v}$ shows the relation of the ore particle diameter to that of the useful component grain impregnated.
These transformations result in a table similar to the initial one but mass fractions are shifted in favor of much finer fractions.

Classification

Then, the ground mass flow goes to classification. Classification is the process of separation of mass flow into 2 parts, according to the particles size.
Classification is characterized by the parameters denoting the probability of classification of particles of a definite size and density into fine or course part of mass flow, with these parameters being specified in a special table. Multiplying the cells of the table corresponding to the mass flow after grinding, and the cells of the classifier characteristic table, it is possible to generate tables for each part of the mass flow where the coarse part will be transported to additional grinding and the finer part – to separation.

Separation

Separation is a process similar to that of classification, i.e., the mass flow is separated into 2 parts. However, unlike classification, separation is based on the useful component content in the mass flow.
To separate mass flow into tailings, concentrate and middlings use is usually made of 2 processes of separation with different characteristics of separators. The first process separates tailings from ore, the second process separates concentrate from the mass flow part left. As to produce concentrate becomes possible only when the size of ore lumps is close to that of the useful component grain impregnated, the initial stages of separation may include only the first process of separation (tailings separation). The middlings produced at the current separation stage is transported to the next stage of separation.
In terms of modeling, knowing the coefficients of the table of mass flow after grinding, and the characteristic tables of classifier and separators, it is possible to calculate the coefficients of the tables for mass flow - tailings, concentrate and middlings:
$\gamma_{ij}^{cl} = \gamma_{ij} * \epsilon_{ij}^{cl}$ are the coefficients of the table of mass flow entering the separator after classification, where γ_{ij} and ϵ_{ij}^{cl} are the coefficients from the table corresponding to the mass flow after grinding and the classifier characteristic table respectively.
$\gamma_{ij}^{t} = \gamma_{ij}^{cl} * (1 - \epsilon_{ij}^{st})$ are the coefficients of the table of tailings after the first process of separation, where ϵ_{ij}^{st} are the coefficients from the characteristic table of the first separator.

The coefficients of the intermediate table of mass flow after tailings separation will be calculated by the formula: $\gamma_{ij}^{cl} * \epsilon_{ij}^{st}$. If the second process of separation is absent, this table becomes the table of the middlings of the current separation stage, otherwise these coefficients will be used to form the tables of concentrate and the middlings of the current separation stage:
$\gamma_{ij}^{c} = \gamma_{ij}^{cl} * \epsilon_{ij}^{st} * \epsilon_{ij}^{sc}$ are the coefficients of the table of concentrate flow where ϵ_{ij}^{sc} are the coefficients from the characteristic table of the second separator.
$\gamma_{ij}^{p} = \gamma_{ij}^{cl} * \epsilon_{ij}^{st} * (1 - \epsilon_{ij}^{sc})$ are the coefficients of the table of the middlings flow of the current separation stage.
As the coefficients of the table of mass flow after grinding depend on the energy spent for grinding, the coefficients of the tables of all the three, produced after separation, flows (tailings, concentrate and middlings) depend on the energy spent for grinding, which in turn is proportional to its volume.
Now it is possible to explicitly write the constraints for the admissible content of the useful component in tailings and concentrate.
The useful component content in the product separated as the tailings:

$$\beta^t = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}\gamma_{ij}^{t} * \beta_i}{\sum_{i=1}^{n}\sum_{j=1}^{m}\gamma_{ij}^{t}} = \\ = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}\gamma_{ij}^{cl} * (1 - \epsilon_{ij}^{st}) * \beta_i}{\gamma^t} \leq P^t$$

γ^t is the tailings output (mass fraction).
P^t is the admissible, according to the norm, content of useful component in tailings.
The useful component content in the product separated as concentrate:

$$\beta^c = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}\gamma_{ij}^{c} * \beta_i}{\sum_{i=1}^{n}\sum_{j=1}^{m}\gamma_{ij}^{c}} = \\ = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}\gamma_{ij}^{cl} * \epsilon_{ij}^{st} * \epsilon_{ij}^{sc} * \beta_i}{\gamma^c} \geq P^c$$

γ^c is the separation concentrate output (mass fraction), $\gamma^t + \gamma^c = 1$.
P^c is the admissible, according to the norm, useful component content of in the concentrate.

CONCLUSION

The analysis of the problem of structural synthesis for effective schemes of multistadial technological processes, which are characterized by the presence of heterogeneous constraints, has shown that the solution of this problem can be reached using the constraints satisfaction methods. The paper presents the initial phase of studies into the development of an information technology of structural synthesis of multistadial flowsheets. Chosen as a "test object" for synthesis is multistadial flowsheets. The problem of structural synthesis for schemes of multistadial flowsheets, with the presence of a complex of heterogeneous performance criteria and constraints is reduced to a set of Constraint Optimization Problems, which allows use of an extensive methods and heuristics which are developed in the constraints satisfaction theory. The application of such an approach makes it possible to jointly process both quantitative and qualitative constraints in the domain.

ACKNOWLEDGEMENTS

The reported study was funded by RFBR, project numbers 18-07-00615-a and 19-07-00359-a.

REFERENCES

Alpatov Y., 2019. Structural-parametric synthesis of multiply-connected control systems (In Russian). Lan'.

Apt K. and Wallace M., 2006. Constraint logic programming using ECLiPSe. Cambridge University Press.

Brochot S., 2002. USIM 3.0: New Features for a global Approach in Mineral Processing Design. In *Proceedings APCOM 2002 Conference*. 477–494.

Buscemi M. and Montanari U., 2008. A survey of constraint-based programming paradigms. *Computer Science Review*, 2, 137–141.

Herbst J., 2004. A microscale look at tumbling mill scale-up using high fidelity simulation. *International Journal of Mineral Processing*, 74, 299–306.

Kreter S.; Rieck J.; and Zimmermann J., 2016. Models and solution procedures for the resource-constrained project scheduling problem with general temporal constraints and calendars. *European Journal of Operational Research*, 251(2), 387–403.

Margaux N.; Artigues C.; and Lopez P., 2017. Cumulative scheduling with variable task profiles and concave piecewise linear processing rate functions. *Constraints*, 22(4), 530–547.

Morrison R. and Richardson J., 2002. JKSimMet: A simulator for analysis, optimisation and design of comminution circuits. In *Symposium, Mineral processing plant design, practice, and control*. Society for Mining, Metallurgy, and Exploration, vol. 1, 442–460.

Rossi F.; van Beek P.; and Walsh T., 2006. Handbook of Constraint Programming. Elsevier.

Russel S. and Norvig P., 2010. Artificial Intelligence: A Modern Approach. Prentice Hall, 3 ed.

Improvement of Commercial Activities Through a better Identification of Exporting Companies using Unstructured Data

Kenneth Van den Bergh, Dries Van Nieuwenhuyse, Younes Ghammad, Andres Van Rompaey, Lode Vermeersch
E-mail: kennethvdbergh@gmail.com

KEYWORDS

Classification, unstructured data, export propensity, similarities, corporate websites, trade credit insurance.

ABSTRACT

This paper is a use case study and aims to investigate the potential added value of unstructured data available on the world wide web for the commercial activity and risk evaluation in a trade credit insurance company.
Identification of potential clients, prospects, is a key challenge for every company. In credit insurance we insure the risk of non payment in a transaction between a seller and a buyer. Therefore finding new clients means finding these sellers that have insurable transactions that are perceived as being risky enough to take an insurance.
The risk for the credit insurer is on the buyer. More often than not these buyers are active in countries where structured data such as financial statements are hard to find.
More than ever companies have websites and use social media channels leaving a lot of digital traces and potential useful information on the web. Although this information might lack objectivity and standardization, when handled in a smart way it can prove to be very useful due to the fact it gives a very recent and qualititave view compared to for example financial statements that give a quantitave view on the performance of last year.
As an experiment we decided to look into the potential of website data to increase the effectivenss of our commercial activity in Belgium. We started with a sample of all active Belgian companies with more than 10 FTE and retrieved the words on the first 50 pages of their website. Using seed words and feature selection we were able to structure the information. Unsupervised learning techniques such as hierarchical clustering and supervised learning techniques such as Boosted trees were used to identify similar companies and model for the propensity of being an exporter.
The results of the first experiment look promising and have already proven to be useful to identify the clients to approach in marketing campaigns and prospect conferences.

INTRODUCTION

A credit insurance company helps companies to manage their trade credit risks by providing trade credit insurance. The goal is to facilitate global trade by offering protection against the risk of nonpayment to companies trading with one another.

One of the bigger challenges for a credit insurer is to evaluate the risk in these countries where data and more importantly structured data such as financial statements are scarce. Since credit insurance is a niche product, identification of potential clients is also very challenging. It often results in a tedious exercise to find relevant information on whether or not companies are exporting goods or services.
In order to improve the effectiveness of its commercial activities and to evaluate risk when little structured information is present an experiment was set up to look into the potential added value of unstructured information on the world wide web.
Being used to work with structured data sources such as financial statements, tapping the other 99 % of data available on the world wide web was a huge step to take. However the goal was clear: to become at least as smart using unstructured data as we are with structured data.

STRUCTURED VERSUS UNSTRUCTURED DATA

Financial

Most companies within OECD countries are well organized and publish their financial results on an annual or quarterly base, in a structured format. This information is typically used for credit modelling in the financial sector as it allows to have an idea on the financial performance, which is the outcome of the strategy and operations.

Non-Financial

The idea behind a company website is to illustrate why a potential customer would start doing business with the company. The website can be considered as a representative reflection of the company at hand (Bojnec and Fertö, 2009). Hence, determining the specific words occurring on the website could result in a textual image that can be quantified and used for stastical analysis. This might help obtaining a view on each company in developing countries that mostly lack detailed financial information.

METHODS

How to turn unstructured data into structured information?

Obtaining a view on the business of a company through the words that occur on its website proved useful for analysis (Blazquez and Domenech, 2018). To obtain a word

frequency matrix of companies, accessible domain names of 9715 companies in the Brussels capital region and Flemish Region with the legal status active with the number of employees of the last available year at least 10 FTE were obtained from the National Bank of Belgium (Belfirst, 2019).
Of each active website the most important webpages were scraped by means of web crawling framework Scrapy implemented in Python. The crawler navigates through websites, downloads each webpage related to the website's domain and parses the content in order to get semi-structured data that will be used for modeling the propensity to export. Some company websites contain over thousand pages and the crawler could start looping. To avoid this problem we force the crawler to scrape only the most important pages per domain assuming that the most relevant content will be captured.

The data extracted from the webpages are the following:

- Textual content: all the text encapsulated in HTML tag is extracted with Selenium Web Driver and automatically translated in English via Google Translate API.
- Language: the language used is detected with langdetect which is a Python Language detection library ported from Google's language-detection
- Countries: countries mentioned in a webpage and their number of occurrences within the document are extracted with geotext Python API that extract geographical entities like countries, cities, and region.
- Export keywords: terms related to exportation and the number of occurences is extracted from the web content. Experts have manually set the keywords list.

This resulted in 492910 webpages read that were submitted to the Text Mining module within Statistica software version 13.5.

Feature selection

In text mining it is all about contextualization. Keywords related to export have a significant influence on the probability to export (Blazquez and Domenech, 2014). To guide the selection of the relevant keywords we first scraped the website of a credit insurer assuming that the terminology ought to reflect the nature of export credit insurance. The word frequency matrix was examined and cleaned up manually by an expert panel for relevance. Secondly we scraped the websites of some customers and looked for complementary keywords. We also included several keywords used in leading export related literature (Sinkovics and Sinkovics, 2012; Anderson et al., 2004). The textual descriptions of NACE codes were used to determine relevant commercial activities and sectors.

The Statistica feature selection functionality was applied to determine the F-value of the correlation between the keyword and the occurrence of the word "export" as a target variable. The p-value of the F-test determined the selection of the keywords for further analysis.

The final bag of words was generated through an iterative accordion-like process of extending up bags of words introducing new words to be found and next squeezing down again by only selecting correlating words with the target variable. Special keywords were looked for and counted e.g. sectors, activities and countries in addition to specific business related terminology. This resulted in 1680 keywords that were used for analysis.

The feature selection method retained 206 indicators for export: 20 sectors, 20 activities, 60 countries, 100 business terms. Six aggregate values were calculated as the total number of countries mentioned, the total number of unique countries, the total number of business terms, the total number of unique business terms, sectors, activities.

Statistical analysis

In order to facilitate the commercial process we used both supervised and unsupervised learning methods. The supervised method models the probability of having the target variable "export" mentioned on the website, the unsupervised methods identifies similar companies based upon the wording of the corresponding websites.

Unsupervised learning

The structured information obtained from the text mining approach yields a word frequency matrix for each company. This matrix serve as input for the calculation of a distance matrix containing Euclidean distances between each website. Sorting the distances allow the identification of similarity in the wording and the ranking of lookalike websites. For operational prospection this proved to be a valid way of working.

Supervised learning

Classification techniques were used to predict the propensity of being an exporter by having export-related keywords on the website of each company. We used CART decision trees, Logistic regression and Boosted Trees to calculate this propensity. Life curves and confusion matrices were calculated to compare the validity of the different models.

Validation

Qualitative validation of the similarity distances was done manually by a commercial subject matter expert who classified the ten closest websites of 30 known exporting companies for their commercial relevance.

RESULTS

Bags of words

The contextualization of the most relevant terms was done in four domains i.e. countries, sectors, activities, business terms. We illustrate the respective bags of words by means of word clouds that are weighed with the average frequencies per keyword in the entire population of companies.

Figure 1. Activities word cloud containing best predictors based upon the p values of the respective F values less than 0.01.

Figure 2. Countries word cloud containing best predictors, based upon the p values of the respective F values less than 0.01.

Figure 3. Sectors word cloud containing best predictors, based upon the p values of the respective F values less than 0.01.

Figure 4. Terms/Products word cloud containing best predictors, based upon the p values of the respective F values less than 0.01.

Similarities

The relevance scores of 300 websites, manually validated for relevance by a subject matter expert for export, activities and similar countries, are given in Table 1.

Table 1. Validation of ten nearest neighbour websites for 30 exporting companies.

Commercial Relevance 10 closest websites	Mean	Std.Dev.
Export (0/1)	8.700000	1.512021
Similar Activities (0/1)	2.600000	1.975715
Similar Countries (0/1)	4.400000	2.268221

The majority of companies appearing in the similarity validation do export or offer services abroad (8.7/10). Europe is the most common market, but quite a few companies are active worldwide. The model has some difficulties with finding similarity between related activities (2.6/10). In general it works better for companies active in manufacturing and food, worse for service based companies, especially if they're active in the ICT sector. The distance based upon similar countries scores mediocre (4.4/10) but displays the highest standard deviation.

Classification

Different Classification models were generated to predict the occurrence of the word "export" on the companies websites. We modelled presence of the term only, not counts. K-Nearest Neighbors, Naïve Bayes Classifiers, Logistic Regression and Boosted Classification Trees were used and compared for predictive power. Data were stratified using all positive observations and 18% of negative observations to obtain a balanced setup ($n_{positive}$=1423, $n_{negative}$=1444). The data was then randomly split into a training and a test set of almost equal size (n_{train}=1438 and n_{test}=1429).

The accuracy of the models was compared through gain charts (Fig. 5) and confusion matrices (Tab. 2).

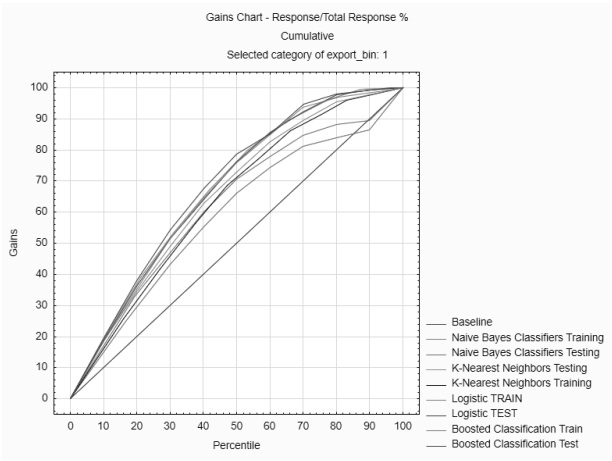


Figure 5. Lift chart comparing 8 classification models.

Table 2. Confusion matrix of the best classification model.

	Confusion matrix of predicted values for export		
	Boosted Classification Test 0 (Predicted)	Boosted Classification Test 1 (Predicted)	Total
Count	524	228	752
Column Percent	81.24%	29.08%	
Row Percent	69.68%	30.32%	
Total Percent	36.67%	15.96%	52.62%
Count	121	556	677
Column Percent	18.76%	70.92%	
Row Percent	17.87%	82.13%	
Total Percent	8.47%	38.91%	47.38%
Count	0	0	0
Column Percent	0.00%	0.00%	
Row Percent			
Total Percent	0.00%	0.00%	0.00%
Count	645	784	1429
Total Percent	45.14%	54.86%	

Boosted Trees Propensity Model

The Boosted Trees models performed best with an overall accuracy of 75.58% correct classifications with a balanced setup for the test dataset. The optimal number of trees was 59 with maximum tree size of 3. The algorithm for Boosting Trees evolved from the application of boosting methods to regression trees. The general idea is to compute a sequence of (very) simple trees, where each successive tree is built for the prediction residuals of the preceding tree. This method will build binary trees, i.e., partition the data into two samples at each split node. Now suppose that you were to limit the complexities of the trees to 3 nodes only: a root node and two child nodes, i.e., a single split. At each step of the boosting trees algorithm, a simple (best) partitioning of the data is determined, and the deviations of the observed values from the respective means (residuals for each partition) are computed. The next 3-node tree will then be fitted to those residuals, to find another partition that will further reduce the residual (error) variance for the data, given the preceding sequence of trees.

Relevance of the insights

Identifying valuable prospects for foreign trade and export credit insurance proved to be difficult using available economic data. We studied the usefulness of text mining and analytics to capture softer information. Relevant keywords as found on the company website allowed us to calculate nearest neighbor websites that yielded a relevance score of 8.7/10 for export features. Classification models yielded Boosted Trees models that performed best with an overall accuracy of 75.58% correct classifications with a balanced setup for the test dataset.

CONCLUSION

Some company features are only visible between the lines and are not as such available. Non-structured data as found on corporate websites proves to be useful in identifying similar websites and hence companies. Classification models return export propensity scores based on the website word use. Hence commercial activities can have a dual approach i.e. starting from an existing customer and identifying similar companies for warm prospection or starting with probabilities of being an exporter for cold prospection. This makes commercial activities more effective.

REFERENCES

Anderson, S; Gabrielsson, J; Wictor, I. 2004. International activities in Small Firms: Examining Factors Influencing the Internationalization and Export Growth of Small Firms, *Canadian Journal of Administrative Sciences* 21 (1) 22-34.

Bel-first, 2019. Update number 236 Software version 70.00 data update 23/11/2017 (n° 23600).

Blazquez, D.; and Domenech, J. 2014. Inferring export orientation from corporate websites, *Applied Economics Letters* 21(7): 509–512.

Blazquez D.; and J. Domenech. 2018. Web data mining for monitoring business export orientation. *Technological and Economic Development of Economy* 24(2): 406–428.

Bojnec, Š.; Fertö, I. 2009. Impact of the Internet on manufacturing trade, *Journal of Computer Information Systems* 50(1): 124–132.

Sinkovics, N.; Sinkovics, R. R.; Jean R.-J. "B." 2013. The internet as an alternative path to internationalization?, *International Marketing Review* 30(2): 130–155.

AUTHOR BIOGRAPHIES

KENNETH VAN DEN BERGH was born in 1986, Belgium and graduated as bio engineer at the University of Ghent in 2009. After 2 years of research in an academic setting at PhD level, he worked for 6 years at KBC Bank and Insurance as a branch manager before joining Credendo

in 2016, where he took up the position of Business Intelligence and Analytics Manager.
kennethvdbergh@gmail.com
Rue Montoyer 3, B-1000 Brussels

DRIES VAN NIEUWENHUYSE was born in 1965, Belgium and graduated as biologist and as computer scientist at the University of Ghent respectively in 1987 and 1989. He took up different Business Intelligence and Analytical positions in a variety of banks and insurance companies before joining Credendo in 2016, where he took up the position of data scientist and statistician.
info@atheneconsult.be
Rue Montoyer 3, B-1000 Brussels

YOUNES GHAMMAD was born in 1990, France and graduated as data scientist at the University of Lille in 2015. He was research assistant in Digital Libraries at TU Braunschweig. After 2 years of research, he worked for AG Insurance as data scientist before joining Credendo in 2018, where he took up the position of data scientist.
Y.Ghammad@credendo.com
Rue Montoyer 3, B-1000 Brussels

LODE VERMEERSCH was born in 1964, Belgium and graduated as civil engineer at the University of Ghent in 1987. He obtained his PhD on identifying ill-defined systems with neural networks in 1998 at the Technical University of Delft. After an international career in banking at KBC Bank and Insurance in Prague, Shanghai and London he joined Credendo in 2013, where he took up the position of Chief Information Officer and is now also Chief Strategy and Innovation Officer.
L.Vermeersch @credendo.com
Rue Montoyer 3, B-1000 Brussels

KSU-STEM Approach and its Possible Algorithmic Revisions verified in a Portfolio Selection Process

Adam Borovička
Department of Econometrics
University of Economics, Prague
W. Churchill Sq. 4, Prague
Czech Republic
E-mail: adam.borovicka@vse.cz

KEYWORDS
Fuzzy, KSU-STEM, multiple objective, open unit trust, portfolio selection.

ABSTRACT

Topic of the article is inspired by usual situation on the capital market – investment portfolio selection. Besides the standardly used methods and approaches, non-traditional multiple objective programming methods can effectively help to make an investment decision. More suitable for this purpose are the methods with an interactive procedure through the portfolio may gradually adapted to the invertor's preferences. Such a method is STEM, or more suitable improved version KSU-STEM. Although this method is an appropriate tool for a portfolio making, it also shows partial algorithmic weaknesses or methodical aspects to think about. Firstly, an unnecessarily negative principle to determination of the basal value of the objectives is modified. Further, the fuzzy goals are specified which leads to a reformulation of the revealed defuzzified multi-objective model. Finally, the imperfect re-setting the weights (importance) of unsatisfactory objectives is revealed. Thus, the alternative approaches are proposed. The interventions to the algorithm are empirically verified through a real-life selection of a portfolio of the open unit trusts offered by Česká spořitelna traded on the Czech capital market. A real-life application shows a significant supporting power of the revised multiple objective programming approach KSU-STEM in a portfolio making process.

INTRODUCTION

Decision making can be very difficult process. Therefore, we should have some supporting tools to make this process easier. Example of usually a complex decision making process is an investment portfolio selection that is facing an increasing number of people trying to valorize their free funds.

To make a portfolio, the intuitive approach, sometimes supported by a knowledge of basic quantitative characteristics of the investment (return, risk, etc.) is often applied. Then the portfolio is especially made on the basis of human intuition, or mood. Such a decision can be supported by the psychological analysis developed by Le Bon (1896). This approach is predominantly qualitative. Portfolio can be also made based on the fundamental (Graham and Dodd, 2008) or technical (Murphy, 1999) analysis. Then we receive a value of some quantitative indicator or "graphical" information based on a historical development of the asset prices. Portfolio selection based only on some qualitative/quantitative information or one-criterion perspective is too simplifying. Moreover, no mentioned approach doesn't enable to make a portfolio, or to explicitly determine shares of the particular assets. They provide only some information on selected investment instruments supporting the following steps leading to a portfolio composition.

Therefore, I propose to use a decision making theory often wrongly neglected in this area. Of course, a well-known Markowitz optimization model (Markowitz, 1959) is sometimes applied. However, this model takes into account only one, or two characteristics (return and risk). Multiple objective programming methods can make a portfolio under qualitative and quantitative input information on the investment instruments or investor's preferences. These multi-criteria approaches are therefore able to provide more complex and representative result. Moreover, they can also work with outputs of the aforementioned analyses.

In my opinion, the most suitable multiple objective programming methods are the interactive approaches, or the methods with a continuous information from the DM. These methods allow a subsequent modification of the actual solution (portfolio) according to the DM's (investor) preferences. Such an approach helps to find a portfolio to the investor's satisfaction. Interactive multiple objective programming methods were developed from 1970s. The first methods were Step Method (STEP) designed by Benayoun et al. (1971) or GDF method proposed by Geoffrion et al. (1972). In 1980s, a stochastic form of the interactive multi-objective methods was introduced (e.g. Teghem et al., 1983). In the same time, a fuzzy form was also proposed (e.g. Sakawa, 1993). Many of fuzzy interactive methods work with α-cut, e.g. Liang (2006). To complete a brief overview of the interactive procedure, interactive goal programming methods have been also designed, e.g. Leung (1987).

Selecting a suitable method is affected by a particular decision making situation, or the DM's abilities. In our case, a portfolio from open unit trusts traded on the Czech capital market by Česká spořitelna is being made. To make a satisfactory investment decision, the following requirements should be fulfilled. The algorithm should not require any additional information difficult to determine by DM-investor (goal, threshold, α-cut, distribution of return, explicit relaxation substitution among objective values, etc.). The importance of the objectives must be adjustable by the investor. Strictly determined relaxation of some objectives

should be acceptable. The algorithm should be user-friendly for its wider and easy applicability. Method fulfilling all these assumptions can solve the investment problem satisfactorily.

An improved form of STEM, KSU-STEM, is shown as a good candidate. However, the algorithm is not perfect. In other words, I see current aspects of the algorithm for reflection or improvement. The first partial question is about a determination of the basal (worst) possible value of the objectives. Unnecessary pessimistic approach is modified. Secondly, the revealed original multi-objective mathematical model is transformed to the one-objective form by introducing a fuzzy goal principle. Such a concept is an effective alternative to the current formulation. Finally, a recalculation of weights of the unsatisfactory objectives within the interactive procedure is evaluated as redundant. This process can be replaced by the simpler ways. The approaches with preserved original weights or with an integration of the weights to the newly determined fuzzy goals are proposed.

The primary aim of this article is to select a suitable method to make an investment portfolio. The main aim is to revise its algorithm to solve the problem as satisfactory as possible. Suitability of these revisions and recommendations is tested on a real-life making the portfolio of open unit trusts offered by Česká spořitelna. More general mission of this paper is to demonstrate a significant application power of (interactive) multiple objective programming methods in the portfolio selection problems in order to use them more in this area.

The structure of the article has the following form. After the introduction, the investment decision making situation is outlined. The next section discusses the interactive multiple objective programming methods. Then, the algorithm of the selected KSU-STEM method is described. Subsequently, the algorithm is revised using demonstrative investment examples. Finally, a portfolio of open unit trusts is made by a revised KSU-STEM. In conclusion, the article is summarized and some ideas for future research are outlined.

INVESTMENT DECISION MAKING SITUATION

Recently, the investing in open unit trusts becomes more and more popular in the Czech Republic. As a long-term client of Česká spořitelna, I have a personal experience with the open unit trusts offered by this bank institution. Another reason for a selection of just these unit trusts is the fact that Česká spořitelna is one of the biggest 'players' on the Czech capital market of a collective investment from the perspective of the funds' property. This company offers mainly mixed, bond and equity open unit trusts.

It is no doubt that the main characteristics of the investment are return and risk. The investment in open unit trusts is also burdened by several fees. In our analysis, the cost is represented by only initial charge that reduces the invested amount. All other fees (management, license, etc.) are projected to the fund property, or to the fund's return. Other criteria can be a locality of the fund investments, traded currency, style of the fund manager or mood on the capital market. All these aspects can be rather taken into account in a 'preselection' phase when potentially suitable funds for a portfolio are chosen. Then 7 mixed, 9 bond and 12 equity funds are included as we can see with all characteristics in the following table (Table 1).

Table 1: Open unit trusts and their characteristics

Fund	Return	Risk	Cost	Fund	Return	Risk	Cost
Controlled Returns Fund	-0.083	3	1	ESPA Czech Corporate Bond F.	-0.031	2	0.5
Stock Mix	0.307	4	3	ESPA Bond Europe High Yield	0.110	3	3.5
Dynamic Mix	0.194	4	1.5	Sporotrend	0.159	6	3
Balanced Mix	0.175	3	1.5	Global Stocks	0.604	5	3
Conservative Mix	0.089	3	1	Top Stocks	0.715	6	3
Life Cycle Fund 2020	0.062	3	1	ESPA Stock Russia	0.606	7	4
Life Cycle Fund 2030	0.197	4	1.5	ESPA Stock Japan	0.572	6	4
Sporoinvest	-0.018	1	0.3	ESPA Stock Istanbul	-0.828	7	4
Sporobond	0.091	2	1	ESPA Stock Global Emer. Mar.	0.358	6	4
Trendbond	-0.193	4	1	ESPA Stock Global	0.702	5	4
Corporate Bond	0.070	3	1	ESPA Stock Europe Property	0.344	6	4
High Yield Bond	0.194	3	1	ESPA Stock Europe Emerging	0.067	6	4
ESPA Portfolio Bond Europe	0.090	3	3.5	ESPA Stock Europe	0.296	6	5
ESPA Czech Govern. Bond F.	-0.036	2	2	ESPA Stock BIOTEC	1.133	7	5

The Czech clients are rather conservative, saving a longer-time horizon. Then the return (in percentage) is calculated as monthly average from the period July 2013 - July 2019 that can reflect an actual longer-time price development. The risk is measured by a standard (for unit trusts) indicator SRRI which is calculated through a standard deviation(s) of returns over the last five years, see more Fund Glossary (2019). Its value is from an integer interval $\langle 1,7 \rangle$. The cost is represented by the initial charge as a percentage of the invested amount. Prices and SRRI are taken from the web of Investment Centre (2019). Returns are calculated in MS Excel.

As mentioned above, the investment strategy is determined longer-time as the most typical strategy on the Czech market with open unit trusts. It usually represents saving money for retirement age or generally for a consumption in more distant future. Purpose of the investment stimulates a conservative approach. Such an investor is more afraid of the investment loss. S/he is willing to lose some part of the return to maintain a lower level of risk. Cost are also considered. However, the initial charge is not so important than other two characteristics. To weaken a potential loss the share of equity funds will be limited. For easier

management, the portfolio should contain a reasonable number of assets (3-5 funds).

DECISION MAKING THEORY APPROACHES FOR A PORTFOLIO SELECTION

A potential of decision making theory principles in a portfolio selection is considerable. Therefore, this chapter focuses on the interactive multiple objective programming methods.

Review of interactive multiple objective programming methods

Interactive multiple objective methods have been developing since the 1970s. According to Fiala (2013), one of the classification aspects can be a character of the expressed trade-offs – explicit or implicit. One of the oldest methods using explicitly expressed trade-off is called GDF (Geoffrion et al., 1972). Explicit trade-off requires an exact quantification of the acceptable relaxation for a couple of the objectives. Some methods (unlike GDF) require this information from the DM, or offer these substitution rates to assess by the DM, as in (e.g.) Zionts-Wallenius method (Zionts and Wallenius, 1976). Such an active role can be too demanding for the DM. Other methods based on this principle are processed by Miettinen (1999). More friendly approach in this aspect is represented by the methods using implicit trade-offs. Then an exact substitution of the objective values is not required. Representative could be older STEP method (STEM) (Benayoun et al., 1971), or younger modification KSU-STEM (Lai and Hwang, 1996), both using a minimization of a distance from the ideal solution. Development has also undergone methods based on the reference point (Wierzbicki et al., 2000). The aspiration and reservation levels for all objectives must be determined by the DM. This often difficult task tries to be simplified with a "starter" neutral solution. Another method can be NIMBUS based on the 'soft' classification of the objectives into up to five classes (see more Miettinen and Mäkelä, 1995). Some approaches even use a regression within the interactive procedure (e.g. Greco et al., 2008). Interactive multiple objective methods can also take into account stochastic elements, see overview in (Adeyefa and Luhandjula, 2011), or fuzzy elements (e.g. Liang, 2006) to capture imprecise or uncertain information. Goal programming models with an interactive revision of the goals have been also proposed (e.g. Leung, 1987).

Selection of a suitable method is namely predetermined by solved decision making problem – investment portfolio making. This process should be manageable for a wider range of users (investors). Most of them are laics. Making, or interactive revision of the portfolio should be instructional. Many methods are burdened by the requirement for additional information from the DM as mentioned above. It can mean determination of objective goals (goal programming), thresholds (fuzzy approach), distribution of stochastic elements (stochastic programming), explicit trade-offs, etc. Such methods are not suitable for our decision making problem. The method should accept the weights of objective functions because this way of expression of their importance is very friendly thanks to a few easily applicable supporting tools. Finally, input data or relaxation of the objectives should be possible to set in the easiest way, i.e. in the strict form.

Under these assumptions, many methods are useless (methods working with explicit trade-offs, goals, reference point, fuzzy or stochastic elements, etc.). Some methods are based on very difficult algorithm (e.g. Zionts-Wallenius method). Some of them are not even able to work with the weights determined by the DM (e.g. STEM). After such a reduction, the KSU-STEM method seems to be a relevant candidate. Its algorithm is comprehensible. The strict weights determined by the DM is accepted. Only implicit trade-offs are required. On the other side, this algorithm is not perfect. In other words, some fragments of the algorithm stimulate subject thinking.

KSU-STEM

Firstly, KSU-STEM is compared with STEM. Its main improvements are accented. Then the algorithm is described.

KSU-STEM vs. STEM

1) The first KSU-STEM positive is a possibility of determination of the weights of objectives by the DM. On the other side, this fact can also be a disadvantage for the DM who are not able to determine the weights. As there are many supportive tools for weights estimation, this will be a minority of cases.
2) KSU-STEM works better with a combination of minimizing and maximizing objective functions. Any transformation of the objective character is not required.
3) The ideal value is determined via the same approach. However, KSU-STEM determines the basal value at the unnecessarily pessimistic level. Unlike the first two cases, this aspect can be considered as a minor drawback of KSU-STEM.
4) On the other side, these two extreme values of the objective functions are artificially used to representative normalization of their values. STEM uses only the ideal value in this process.
5) Both methods zero the weights of satisfactory objectives. KSU-STEM, compared to STEM, recalculates the weights of other objectives which is actually necessary action (see more below).

KSU-STEM algorithm

The algorithm (with own customized notation) is described step by step according to Lai and Hwang (1996).

Step 1: Define k objective functions (objectives) $f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x})$, where $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ is a vector of n variables. Distinguish the set of indices of the minimizing, or maximizing objective functions denoted as J_{min}, or J_{max}. The importance of the criteria is expressed by the vector of weights $\mathbf{w} = (w_1, w_2, \ldots, w_k)^T$, where $w_j, j = 1, 2, ..., k$, is the weight of the j-th objective function.

Step 2: Minimum and maximum of each j-th objective function are determined on the set of feasible solutions X (containing all conditions of the solved problem) by finding a solution of the following models

$$\begin{array}{ll} \min\ f_j(\mathbf{x}) & \max\ f_j(\mathbf{x}) \\ \mathbf{x} \in X & \mathbf{x} \in X \end{array}. \qquad (1)$$

Solutions of these models represent ideal f_j^I (the lowest, or highest value for minimizing, or maximizing objective function) and basal value f_j^B (the highest, or lowest value for minimizing, or maximizing objective function) of each *j*-th objective function.

Step 3: The mathematical model minimizing a maximal weighted relative (standardized) deviation from the ideal solution (value) is formulated as follows

$$\begin{aligned} &\min\ \alpha \\ &w_j \frac{f_j(\mathbf{x}) - f_j^I}{f_j^B - f_j^I} \le \alpha \qquad j \in J_{\min} \\ &w_j \frac{f_j^I - f_j(\mathbf{x})}{f_j^I - f_j^B} \le \alpha \qquad j \in J_{max}\,. \qquad (2) \\ &\mathbf{x} \in X \\ &0 \le \alpha \le 1 \end{aligned}$$

The optimal solution is denoted as $\mathbf{x}^*$ with the values of the objective functions $f_1(\mathbf{x}^*), f_2(\mathbf{x}^*), \ldots, f_k(\mathbf{x}^*)$. If the values of all objective functions are acceptable by the DM, the compromise solution is found. If the values of all objective function are nonacceptable, the algorithm is also terminated. It is not possible to simultaneously improve all objective values. If the values of some objectives are acceptable and some not, then the interactive procedure can be started to reveal the compromise solution.

Step 4: To improve the values of unsatisfactory objectives, the DM must relax at least one satisfactory objective. Let us denote the set J_{min}^R, or J_{max}^R containing the indices of the minimizing, or maximizing objective functions with satisfactory value. Then the DM provides $\Delta f_j(\mathbf{x}^*), j \in J_{min}^R$, or $\Delta f_j(\mathbf{x}^*), j \in J_{max}^R$, as the amount of acceptable relaxation for the *j*-th minimizing, or maximizing objective function. Then specify the set X^1 including the following conditions

$$\begin{aligned} &f_j(\mathbf{x}) \le f_j(\mathbf{x}^*) \quad j \in J_{min} - J_{min}^R, \quad f_j(\mathbf{x}) \ge f_j(\mathbf{x}^*) \quad j \in J_{max} - J_{max}^R \\ &f_j(\mathbf{x}) \le f_j(\mathbf{x}^*) + \Delta f_j(\mathbf{x}^*) \quad j \in J_{min}^R \\ &f_j(\mathbf{x}) \ge f_j(\mathbf{x}^*) - \Delta f_j(\mathbf{x}^*) \quad j \in J_{max}^R \end{aligned} \qquad (3)$$

Finally, the weights are reformulated to the following form

$$\begin{aligned} &w_j' = 0 \qquad j \in J_{min}^R \cup J_{max}^R \\ &w_j' = \frac{w_j}{\sum_{\substack{j=1 \\ j \notin J_{min}^R \cup J_{max}^R}}^{k} w_j} \qquad j \in J_{min} - J_{min}^R, j \in J_{max} - J_{max}^R\,. \end{aligned}$$

The following modified model (2) is solved

$$\begin{aligned} &\min\ \alpha \\ &w_j' \frac{f_j(\mathbf{x}) - f_j^I}{f_j^B - f_j^I} \le \alpha \qquad j \in J_{min} - J_{min}^R \\ &w_j' \frac{f_j^I - f_j(\mathbf{x})}{f_j^I - f_j^B} \le \alpha \qquad j \in J_{max} - J_{max}^R\,. \qquad (4) \\ &\mathbf{x} \in X \cup X^1 \\ &0 \le \alpha \le 1 \end{aligned}$$

Optimal solution of the model (4) is another compromise solution. If this solution is not still acceptable by the DM, the procedure of changes is repeated. Otherwise, the compromise solution is found, and the algorithm is terminated.

REVISIONS OF THE KSU-STEM ALGORITHM

The main part of the article introduces the possible revisions of the KSU-STEM algorithm which further enhance the already great application (and theoretical) power. Let's proceed step by step of the algorithm.

Step 2: Basal and ideal value of the objectives

The way for a determination of the objective ideal value is meaningful. It is a minimum $f_j(\mathbf{x}_\mathbf{j}^*) = f_j^I, j \in J_{min}$, or maximum $f_j(\mathbf{x}_\mathbf{j}^*) = f_j^I, j \in J_{max}$ of the *j*-th minimizing, or maximizing objective function on the set of feasible solutions. Thus, the following holds for the *j*-th minimizing, or maximizing objective function

$$\begin{array}{ll} \mathbf{x}_\mathbf{j}^* = \arg\min f_j(\mathbf{x}) & \mathbf{x}_\mathbf{j}^* = \arg\max f_j(\mathbf{x}) \\ \mathbf{x} \in X & \mathbf{x} \in X \end{array}, \text{ or } \quad .$$

Related models are solvable because the set *X* is bounded in the portfolio making problem (at least) due to the 'portfolio' condition (unit sum of the assets' share). Of course, some objective function(s) can be nonlinear. Then a local optimum can be found.

The opposite extremes for all objectives are used for a determination of the basal (worst) value. This concept is unnecessarily pessimistic. More reasonable concept determines the basal value of objective function with the respect to the optimal solution for other objective functions. This concept is actually used in the STEM algorithm. Then the basal value of the *j*-th minimizing, or maximizing objective function is specified as follows

$$\begin{aligned} &f_j^B = \max_{\mathbf{x}_\mathbf{j}^*} \left[f_j(\mathbf{x}_\mathbf{j}^*) \right] \qquad j \in J_{min}, \text{ or} \\ &f_j^B = \min_{\mathbf{x}_\mathbf{j}^*} \left[f_j(\mathbf{x}_\mathbf{j}^*) \right] \qquad j \in J_{max}\,. \qquad (5) \end{aligned}$$

The different approach can be practically illustrated on our investment portfolio making problem. The basal values via KSU-STEM and modified approach are presented in the following table (Table 2).

Table 2: Basal values of the objectives via two approaches

Approach	Return	Risk	Cost
KSU-STEM	-0.3505	4.9	3.95
Modified approach	-0.0015	4.9	3.15

It confirms that the result of KSU-STEM procedure provides more negative result than the modified approach.

Step 3: Fuzzy goal construction

In the third step, one-objective model (2) is solved. This model is based on a minimization of maximal weighted relative deviation from the ideal solution. Though the authors of KSU-STEM don't explicitly mention, this model may be a transformation of the following multi-objective model

$$\begin{aligned} &\min \ f_j(\mathbf{x}) && j \in J_{min} \\ &\max \ f_j(\mathbf{x}) && j \in J_{max} \ . \\ &\mathbf{x} \in X \end{aligned} \tag{6}$$

Used principle is derived from the STEM approach that minimizes a deviation from the ideal solution measured by Tchebychev metric. Then a linearized one-objective model leads to *minmax* optimization problem.

In my opinion, another (alternative) concept for finding a solution of the model (6) should be introduced. This principle is actually opposite to the original form. It represents a positive view – maximization 'something'. In the first step, the fuzzy goals must be specified. Then, the *j*-th minimizing objective function $f_j(\mathbf{x}), j \in J_{min}$, is actually transformed to the fuzzy goal represented by the fuzzy set $\tilde{F}_{f_j(\mathbf{x})}, j \in J_{min}$, with the following membership function

$$\mu_{\tilde{F}_{f_j(\mathbf{x})}}[f_j(\mathbf{x})] = \begin{cases} 1 & f_j(\mathbf{x}) \le f_j^I \\ \frac{f_j^B - f_j(\mathbf{x})}{f_j^B - f_j^I} & f_j^I \le f_j(\mathbf{x}) \le f_j^B \\ 0 & f_j(\mathbf{x}) \ge f_j^B \end{cases} . \tag{7}$$

The *j*-th maximizing objective function is also transformed to the fuzzy goal as the fuzzy set $\tilde{F}_{f_j(\mathbf{x})}, j \in J_{max}$, with the membership function

$$\mu_{\tilde{F}_{f_j(\mathbf{x})}}[f_j(\mathbf{x})] = \begin{cases} 1 & f_j(\mathbf{x}) \ge f_j^I \\ \frac{f_j(\mathbf{x}) - f_j^B}{f_j^I - f_j^B} & f_j^B \le f_j(\mathbf{x}) \le f_j^I \\ 0 & f_j(\mathbf{x}) \le f_j^B \end{cases} . \tag{8}$$

The membership function (7) and (8) can be graphically displayed as follows (Figure 1).

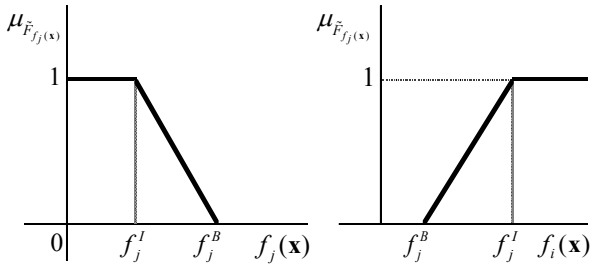


Figure 1: Membership function of minimizing (left) and maximizing (right) fuzzy goal, Source: self-designed

This is actually right and left hand triangular fuzzy number. This type of fuzzy number is selected because its membership function is piecewise linear. Therefore, the work with them is easier. In addition, a shape of the membership function enables to reflect the relationship between the objective value and membership grade well. Thanks to a fuzzy goal principle, the multiple objective model (6) can be transformed to one-objective form as follows

$$\begin{aligned} &\max \ \alpha \\ &(1-w_j)\frac{f_j^B - f_j(\mathbf{x})}{f_j^B - f_j^I} \ge \alpha \qquad j \in J_{min} \\ &(1-w_j)\frac{f_j(\mathbf{x}) - f_j^B}{f_j^I - f_j^B} \ge \alpha \qquad j \in J_{max} , \\ &\mathbf{x} \in X \\ &0 \le \alpha \le 1 \end{aligned} \tag{9}$$

where α is an intersection of all fuzzy sets in the model (9) representing the (weighted) membership grade of a solution. Therefore, α is maximized. A real (not weighted) membership grade of solution is as follows

$$\min_{j \in J_{min} \cup J_{max}} \left(\frac{f_j^B - f_j(\mathbf{x})}{f_j^B - f_j^I}, \frac{f_j(\mathbf{x}) - f_j^B}{f_j^I - f_j^B} \right).$$

Let's get back to the weights in the model (9). As we can see the position of the weights is different compared to the model (2) to reflect the importance of the objectives correctly. Of course, the same principle of a weight position is used in the model (11) within an interactive procedure. Other way of an importance expression is to divide the left side of the conditions (related to the fuzzy goal) by the weight. It is an alternative approach that however requires to an elimination of the condition for the α values because α could then be greater than 1. On the contrary, in the (revised) model (9) the condition $0 \le \alpha \le 1$ is actually necessary because the α value is always in this interval. The weights can be estimated by any suitable existing method, e.g. a well-known user-friendly scoring method (Fiala, 2013).

Step 4: Weight recalculation

Another revision of the KSU-STEM algorithm deals with a recalculation of the weights within the interactive procedure. In KSU-STEM and STEM algorithms, the weights of satisfactory objectives become zero. The process of recalculation of the weights of unsatisfactory objectives is different. The STEM algorithm leaves the weights in the original form. In KSU-STEM, the weights are recalculated to the standardized form. Thus, the sum of weights is 1.

Special case of single unsatisfactory objective

Firstly, we introduce the special case of one unsatisfactory objective. In the original KSU-STEM algorithm, a unit weight of unsatisfactory objective works well in the model (2). However, the problem occurs in the modified model (9) because zero is on the left side of the condition related to the fuzzy goal. This fact distorts the results. In such a situation, the following modification is proposed. The weight is transformed to the form of 'at least one'. Then one is reduced by the infinitesimal constant as follows

$$w_j = 1 - 10^{-8} \qquad j \in J^S , \tag{10}$$

where $J^S = \{J_{min} - J^R_{min}\} \cup \{J_{max} - J^R_{max}\}$ is the single element set containing the index of one unsatisfactory objective.

Necessity of the weight recalculation?

The main question is about the need to recalculate the weights. Of course, the recalculation does not change the relative relations among the weights. On the other side, the absolute differences are naturally modified. Given again our investment portfolio making problem. After all conditions (see below in more detailed), the initial portfolio has the following form: 40% *Sporoinvest*, 38.58% *Sporobond* and 21.42% *ESPA Stock Biotec*. The monthly return is 0.271%, risk 2.671% and cost 1.577%. Under the risk-averse strategy and knowledge of the extreme values of the objectives, the investor is not satisfied with the level of portfolio risk. To reduce this value, the investor is able to sacrifice some part of return and accept the cost increase. Then the return can be reduced by 0.03 percentage point and the cost can be increased by 0.4 percentage point. Via the revised KSU-STEM, the weight of risk is determined by (10). The weights of other two objectives are zero. Then the portfolio has the following structure: 30% *Sporoinvest*, 40% *Sporobond* and 30% *ESPA Stock Global*. The main portfolio characteristics are: return 0.241 %, risk 2.59%, cost 1.69%. If the STEM concept for a weight 'recalculation' is used (original weight of risk, zero weights of return and cost), the result is the same. Other combinations of the objectives for their value improvement were selected (return and cost, risk and cost), a composition of the portfolios is the same by means of both approaches. The main reason is that the relative relations among the weights of unsatisfactory objectives remain the same.

After all previous revisions, the model for a solution improvement in terms of the interactive procedure is formulated as

$$\begin{aligned} &\max\ \alpha \\ &(1-w_j)\frac{f_j^B - f_j(\mathbf{x})}{f_j^B - f_j^I} \geq \alpha \qquad j \in J_{min} - J^R_{min} \\ &(1-w_j)\frac{f_j(\mathbf{x}) - f_j^B}{f_j^I - f_j^B} \geq \alpha \qquad j \in J_{max} - J^R_{max} \\ &\mathbf{x} \in X \cup X^1 \\ &0 \leq \alpha \leq 1 \end{aligned} \tag{11}$$

The set of conditions X^1 formulated within the interactive procedure should be also discussed. The original formulation (3) is no perfect because it also enables a non-improvement unsatisfactory objective thanks to the following expression

$$\begin{aligned} f_j(\mathbf{x}) &\leq f_j(\mathbf{x}^*) \quad j \in J_{min} - J^R_{min} \\ f_j(\mathbf{x}) &\geq f_j(\mathbf{x}^*) \quad j \in J_{max} - J^R_{max} \end{aligned}.$$

Mainly in the case of a higher number of unsatisfactory objectives, the conditions should be transformed to the sharp inequalities

$$\begin{aligned} f_j(\mathbf{x}) &< f_j(\mathbf{x}^*) \quad j \in J_{min} - J^R_{min} \\ f_j(\mathbf{x}) &> f_j(\mathbf{x}^*) \quad j \in J_{max} - J^R_{max} \end{aligned}. \tag{12}$$

Sharp inequalities (12) may be a problematic for searching a solution of the models. Then the conditions can be transformed via the infinitesimal constant as follows

$$\begin{aligned} f_j(\mathbf{x}) &\leq f_j(\mathbf{x}^*) - 10^{-8} \quad j \in J_{min} - J^R_{min} \\ f_j(\mathbf{x}) &\geq f_j(\mathbf{x}^*) + 10^{-8} \quad j \in J_{max} - J^R_{max} \end{aligned}. \tag{13}$$

This formulation holds only under the assumption that the sets J^R_{min} and J^R_{max} contains the indices of all satisfactory objectives, including cases with the zero relaxation value. Although the authors do not explicitly declare this fact, I consider it reasonable. If not, the weights of satisfactory objective would be also taken into account in the nonzero form. Then the conditions/fuzzy goals with these weights would not be eliminated. The result could be distorted. Similar situation could arise by maintaining the non-zero weights of the satisfactory objectives with a particular acceptable relaxation level. Explicitly expressed preference (by means of the weights, or relevant conditions) about a possible relaxation can inadequately 'muffle' this allowed relaxation against the unsatisfactory objectives. α can be then improperly reduced. Our investment situation confirms this declaration. After a formulation of the additional preferences of a portfolio improvement mentioned above, a new portfolio, obtained as described, has the return at the level of 0.26%, risk 2.6% and cost 1.52%. As we can see, the risk is worse a little (vs. 2.59%) and return is better (vs. 0.241%). Under the additional preferences, this solution is evidently unnecessarily bad.

Thus, the sets J^R_{min} and J^R_{max} must contain all satisfactory objectives. Then the revised model (11) with the set X^1 modified by (13) reflecting the DM's requirements for a solution improvement. The weights of unsatisfactory objectives are explicitly included. The weights (importance) of satisfactory objectives are presented implicitly through the minimum, or maximum restrictive value in the revised set X^1.

Fuzzy goal modification

In terms of the (revised) KSU-STEM interactive procedure, the additional preferences, or the modified importance (weight) of the objectives can be reflected by another way (besides two approaches described above). I propose a modification of the established fuzzy goals for the unsatisfactory objectives. A stronger preference can be taken into account by an improvement of the basal value. It means its decrease, or increase for a minimizing, or maximizing objective. A modification of the fuzzy goals can be graphically displayed as follows (Figure 2).

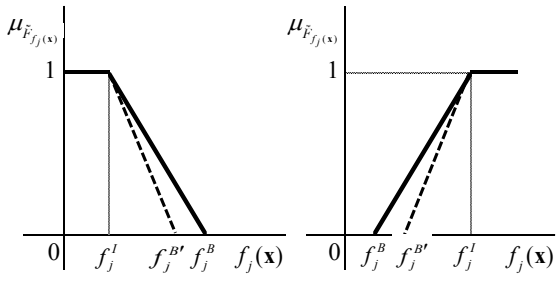


Figure 2: Modification of minimizing (left) and maximizing (right) fuzzy goal, Source: self-designed

It is obvious that a modified basal value pushes to improve value of the *j*-th objective function $f_j(\mathbf{x})$. The main question is how to get a new basal value $f_j^{B'}$. I propose an approach based on the integration of the weight to the fuzzy goal, or (triangular) fuzzy number. Then the following formula holds for the *j*-th minimizing, or maximizing fuzzy goal

$$f_j^{B'} = f_j^B - w_j\left(f_j^B - f_j^I\right) \qquad j \in J_{min} - J_{min}^R, \text{ or}$$
$$f_j^{B'} = f_j^B + w_j\left(f_j^I - f_j^B\right) \qquad j \in J_{max} - J_{max}^R .$$

Then the modified fuzzy goals are in the following form

$$\frac{f_j^{B'} - f_j(\mathbf{x})}{f_j^{B'} - f_j^I} \geq \alpha \qquad j \in J_{min} - J_{min}^R$$
$$\frac{f_j(\mathbf{x}) - f_j^{B'}}{f_j^I - f_j^{B'}} \geq \alpha \qquad j \in J_{max} - J_{max}^R .$$

Other part of the model (11) used in the interactive procedure remain unchanged. If this concept is applied to the aforementioned case of risk reduction, the result (portfolio composition) will be the same. The same situation also occurs for a mentioned combination of improved objectives (return and cost or risk and cost). So, this approach can be comprehended as an alternative to the approach with the explicitly expressed weights of unsatisfactory objectives.

Now we have actually three alternative approaches to taking into account the preferences of an unsatisfactory objective value(s) – the original KSU-STEM with a weight recalculation, revised approach with the original weight of unsatisfactory objective and revised approach of modified fuzzy goals via the integrated weights. The second approach has one advantage compared to two others. It is easier to apply thanks to actually no need to recalculate the weights or basal values. Maybe, a formulation of fuzzy goal is simpler because of the integrated weights. However, the revised approach with the original weight is clearly the most user-friendly.

SELECTING A PORTFOLIO OF OPEN UNIT TRUSTS VIA REVISED KSU-STEM

Finally, a revised version of KSU-STEM is applied to making a portfolio of the open unit trusts offered by Česká spořitelna. As mentioned above, it is a longer-time risk-averse investment. The weights of three determined criteria are calculated via a scoring method: return – 0.35, risk – 0.5, cost – 0.15. For easier portfolio management, the number of open unit trusts is limited by minimum (15%) and maximum (40%) share in the portfolio. Further, the risk-averse investor limits a share of equity funds to 30% because these funds potentially generate the highest loss. All strategy aspects can be determined by a "more advanced" investor or with the assistance of an experienced investment counsel. The same applies to the application of the method itself.

The ideal values of three objectives are determined through (1), where the set of feasible solutions *X* is following

$$X = \left\{ \begin{array}{l} 0.15y_i \leq x_i \leq 0.4y_i \quad i = 1,2,...,28 \\ \sum_{i=17}^{28} x_i \leq 0.3 \\ \sum_{i=1}^{28} x_i = 1 \\ x_i \geq 0,\ y_i \in \{0,1\} \quad i = 1,2,...,28 \end{array} \right\},$$

where $x_i, i = 1,2,...,28,$ is a share of the *i*-th unit trust (indexed in the order from Table 1). Binary variable $y_i, i = 1,2,...,28,$ helps to ensure (within the first conditions) a fixed interval for the share of each *i*-th open unit trust if it is in the portfolio. Second row represents the limit of equity fund share in the portfolio. Further, there is a standard condition for making a portfolio as a whole. Non-negativity and binarity conditions for the relevant variables must not be missed. Now the initial investment portfolio is made by model (9).

$$\max \ \alpha$$
$$(1-0.35)\frac{\sum_{i=1}^{28} r_i x_i - (-0.0015)}{0.522-(-0.0015)} \geq \alpha$$
$$(1-0.50)\frac{4.9-\sum_{i=1}^{28} l_i x_i}{4.9-1.6} \geq \alpha \qquad (1-0.15)\frac{3.15-\sum_{i=1}^{28} c_i x_i}{3.15-0.52} \geq \alpha ,$$
$$x_i \in X \quad i = 1,2,...,28$$
$$0 \leq \alpha \leq 1$$

where $r_i, l_i, c_i, i = 1,2,...,28,$ is return, risk and cost of the *i*-th open unit trust. The (linear) functions $\sum_{i=1}^{28} r_i x_i, \sum_{i=1}^{28} l_i x_i, \sum_{i=1}^{28} c_i x_i$ represent a return, risk and cost of the portfolio. The basal values are calculated according to the formulas (5). All models are solved in the LINGO optimization software.

The portfolio has the following form (as mentioned above): 40% *Sporoinvest*, 38.58% *Sporobond* and 21.42% *ESPA Stock Biotec*. Return is 0.271%, risk 2.671% and cost 1.577%. Not surprisingly the funds with the lowest level of risk Sporobond and Sporoinvest are in the portfolio in (almost) the highest possible share. Selection of ESPA Stock Biotec fund with the highest risk level is affected by its significantly highest return. And yet the yield has also a significant weight. A weighted grade of membership (α) is 0.338. A real grade of membership of the solution is 0.52. The highest grade of membership is, of course, for a risk (0.676) as the most important objective. The lowest membership grade is 0.52 related to the return. Cost as the least important objective has a higher grade of membership than return which might be strange. However, in this case it is natural because the lower-risky open unit trusts usually provide a lower level of cost.

As a risk-averse investor, the risk is comprehended as too high. In order to reduce it, the return can be relaxed by 0.03 percentage point and cost by 0.4 percentage point. These requirements are specified also based on the extreme objective values. Risk cannot be ultimately improved without a return relaxation. To find another solution, the model (11) is solved

$$\begin{aligned} &\max\ \alpha \\ &(1-0.5)\frac{4.9-\sum_{i=1}^{28} l_i x_i}{4.9-1.6} \geq \alpha \\ &x_i \in X \cup X^1 \quad i = 1,2,...,28 \\ &0 \leq \alpha \leq 1 \end{aligned},$$

where the set X^1 is specified in the pattern (3) and (13) as

$$X^1 = \left\{ \begin{array}{l} \sum_{i=1}^{28} r_i x_i \geq 0.271 - 0.03 \\ \sum_{i=1}^{28} l_i x_i \leq 2.671 - 10^{-8} \\ \sum_{i=1}^{28} c_i x_i \leq 1.577 + 0.4 \end{array} \right\}.$$

The portfolio has the following form: 30% *Sporoinvest*, 40% *Sporobond* and 30% *ESPA Stock Global*. The main portfolio characteristics are: return 0.241%, risk 2.59%, cost 1.69%. The riskiest fund is replaced by another equity fund ESPA Stock Global whose risk is two points lower. A grade of membership of the solution is, of course, lower (0.43) due to a return reduction. A membership grade of risk is logically higher (0.7). If the same adjustment were repeated, the portfolio would have the following characteristics: return 0.211%, risk 2.399%, cost 1.519%. Cost is lower because the participants with lower level of risk mostly provide a lower cost. The risk is slowly approaching its ideal value. Its another reduction cannot be too striking. Moreover, the return must significantly fall. Therefore, the investor accepts the current portfolio (40% *Sporoinvest*, 33.38% *Sporobond* and 26.62% *ESPA Stock Global*) with the aforementioned characteristics.

The introduced method accepts any investment strategy, e.g. 'risk-seeking' for potentially higher return. Then the equity funds (generally with the highest risk) would be in a maximum feasible share in the portfolio. This approach could be applied to making a portfolio from another financial investment instruments (stocks, bonds, investment certificates, etc.). However, it can also be used in another field – project management, production systems, etc.

A non-traditional multi-objective approach of decision making theory seems to be very effective tool for a portfolio making. Compared to the one-criterion (fundamental, technical analysis, etc.), or 'human intuition' concepts, it provides a far more complex view to a portfolio selection. Moreover, the investor knows exactly what part of his budget is invested in what investment instrument. Unlike a well-known mean-variance optimization approach, the proposed concept enables to consider a wide range of quantitative, or qualitative investment criteria as well as all investor's preferences. It is evident that an application power of the proposed concept grows with an increasing number of the criteria where the solution is very difficult to predict which is eventually confirmed by the presented investment case.

Thus, a real-life making a portfolio of open unit trusts proves such an application power of the proposed revised KSU-STEM method in the field of the capital market. The interactive procedure is very helpful to obtain a satisfactory portfolio composition. The approach is user-friendly. The interactive procedure is simplified by an easier re-setting the weights. Revised determination of basal objective values avoids finding an unnecessarily pessimistic solution. Sharp conditions (12), or modified (13) also contribute to more effective finding a compromise solution. Then the process is faster which is also an important aspect today.

CONCLUSION

The article deals with a revision of multiple objective programming method KSU-STEM. Inappropriate technique for a determination of the basal objective value is replaced by 'less negative' approach. More effective ways for expressing the importance of the objectives throughout a decision making process are proposed. All algorithmic modifications, or recommendations, lead to more satisfactory applicability which is demonstrated on a real-life selecting portfolio from the open unit trusts traded on the Czech capital market. In general, the article shows a significant usability of the methods of decision making theory in this area where other concepts are applied more often. I think it is a shame!

In the future research, the proposed revisions may be reviewed on other real or simulated data set. The revised KSU-STEM can be also extended by the possibility of expressing often only vague (uncertain) preferences within the interactive procedure (via the fuzzy sets).

REFERENCES

Adeyefa, A.S. and M.K. Luhandjula. 2011. "Multiobjective Stochastic Linear Programming: An Overview." *American Journal of Operations Research* 1, No. 4 (Jul), 203-213.

Benayoun, R.; De Montgolfier, J.; Tergny, J.; and O. Laritchev. 1971. "Linear Programming with Multiple Objective Functions: Step Method (STEM)." *Mathematical Programming* 1, No. 1 (Dec), 366-375.

Fiala, P. 2013. *Models and Methods of Decision Making* (in Czech). Oeconomica, Prague.

Geoffrion, A.M.; Dayer, J.S.; and A. Feinberg. 1972. "An Interactive Approach for Multicriterion Optimization with an Application to the Operation an Academic Department." *Management Science* 19, No. 4 (Dec), 357-367.

Graham, B. and D. Dodd. 2008. *Security Analysis*. McGraw-Hill Education, Columbus.

Greco, S.; Mousseau, V.; and R. Slovinski. 2008. "Ordinal Regression Revisited: Multiple Criteria Ranking with a Set of Additive Value Functions." *European Journal of Operational Research* 191, No. 2 (Dec), 416-436.

Lai, Y.J. and C.L. Hwang. 1996. *Fuzzy Multiple Objective Decision Making: Methods and Applications*. Springer, Berlin.

Le Bon, G. 1896. *The Crowd: Study of the Popular Mind*. Macmillan.

Leung, Y. 1987. "Hierarchical Programming with Fuzzy Objective and Constraints." In *Optimization Models Using Fuzzy Sets and Possibility Theory*, J. Kacprzyk and S.A. Orlovski (Eds.). D. Reidel, Dordrecht, 245-257.

Liang, T.-F. 2006. "Distribution Planning Decisions Using Interactive Fuzzy Multi-Objective Linear Programming." *Fuzzy Sets and Systems* 157, No. 10 (May), 1303-1316.

Markowitz, H.M. 1959. *Portfolio Selection: Efficient Diversification of Investments*. John Wiley & Sons, Inc., New York.

Mietinen, K. 1999. *Nonlinear Multiobjective Optimization*. Kluwer Academic Publishers, Boston.

Mietinen, K. and M.M. Mäkelä. 1995. "Interactive Bundle-Based Method for Nondifferentiable Multiobjective Optimization: NIMBUS." *Optimization* 34, No. 3 (Mar), 231-246.

Murphy, J.J. 1999. *Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications*. New York Institute of Finance, New York.

Sakawa, M. 1993. *Fuzzy Sets and Interactive Multiobjective Optimization*. New York: Plenum Press.

Teghem, J.; Dufrane, D.; Thauvoye, M.; and P. Kunsch. 1986. "STRANGE: An Interactive Method for Multi-Objective Linear Programming under Uncertainty." *European Journal of Operations Research* 26, No. 1 (Jul), 65-82.

Wierzbicki, A.P.; Makowski, M.; and J. Wessels. 2000. *Decision Support Methodology with Environmental Applications*. Kluwer Academic Publishers, Dodrecht.

Zionts, S. and J. Wallenius. 1976. "An Interactive Programming Method for Solving the Multiple Criteria Problem." *Management Science* 22, No. 6 (Feb), 652-663.

WEB REFERENCES

Fund Glossary [online], available at: *http://fundglossary.erste-am.com/srri/*, [cit. 01-07-2019].

Investment Center [online], available at: *https://cz.products.erstegroup.com/Retail/cs/index.phtml*, [cit. 01-07-2019].

ACKNOWLEDGEMENT

The research project was supported by Grant No. IGA F4/66/2019 of the Internal Grant Agency, Faculty of Informatics and Statistics, University of Economics, Prague.

ADAM BOROVIČKA was born in Prague, Czech Republic. He graduated from the University of Economics, Prague in Mathematical Methods in Economics (Ing.). In 2015 he received Ph.D. degree in Econometrics and Operations Research at the University of Economics, Prague. Since 2015 he has been an assistant professor at the Department of Econometrics at the University of Economics and Faculty of Nuclear Sciences and Physical Engineering of CTU in Prague. He primarily deals with the questions of multiple criteria decision making, fuzzy optimization problems and stochastic processes in the field of the capital market.

WAVELET THEORY: FOR ECONOMIC & FINANCIAL CYCLES

Farai F. Mlambo[1] and Igor N. Litvine[2]
Department of Statistics - Nelson Mandela University
email[1]: farai.mlambo@wits.ac.za
email[2]: igor.litvine@mandela.ac.za

KEYWORDS
Cycles, Wavelets, Smoothing, Time Series, Hierarchical Dating

ABSTRACT

In this study, a scalable, smooth and differentiable quarter-circular wavelet basis for the smoothing of business, economic and financial time series is proposed. Using the Hierarchical Cycle Dating methodology proposed by Litvine (2016), a new method for the dating of cycles is proposed, by bringing wavelets, circle geometry and non-linear regression analysis together. The Hierarchical approach, whose disadvantage is that it yields a straight-line cycle dating is used for initial approximations for the Quarter-Circular Wavelet Cycle Dating method proposed in this paper. The models are implemented in Wolfram Mathematica ®, which affords efficient algorithms for symbolic, numerical and functional computation of the proposed methodology. Practical illustrations are given, from South African and other international indices.

INTRODUCTION

Cycles, their nature in existence, their implications on human-kind (directly or otherwise) and the study thereof have (without controversy) sparked some important philosophical debates since the very prehistoric days. Notable contributions by famous, genius Philosophers and Economists such as Kuznets (a Noble Prize winner), Kondratieff, Elliot and many others in itself shows how cycles and their study have been deemed important, through the history of scientific and philosophical inquiry.

Particularly, the explication of Business, Economic and Financial cycles have seen some significant research and policy attention. Nevertheless, most of the currently-used methodologies are either purely empirical in nature (rules of thumb e.g. the Bry and Boschan - BBQ Algorithm), time series based (e.g. isolation of cyclic and other components - Econometric approach) or the so-called Regime-Switching Markov model popularized in Economics by Hamilton. In this work, we develop a Statistical model fit applicable for the dating of cycles.

This study proposes a scalable, smooth and differentiable quarter-circular wavelet basis for the smoothing and/or dating of business, economic and financial cycles. The dating then necessitates the forecasting of the cyclical patterns in the evolution of economic and financial time series. The practical significance of dating and forecasting business and financial cycles cannot be over-emphasized. The use of wavelet decomposition in explaining cycles can be seen as an critical contribution of spectral methods of statistical modeling to finance and economic policy at large.

Being a relatively new method, wavelet analysis has seen some great contribution in geophysical modeling. This study endeavors to widen the use and application of frequency-time decomposition to the economic and financial space. Wavelets are localized in both time and frequency, such that there is no loss of the time resolution. The importance of time resolution in dating of cycles is another motivation behind using wavelets. Moreover, the preservation of time resolution in wavelet analysis is a fundamental strength employed in the dating of cycles.

LITERATURE

According to Aboufadel and Schlicker, Aboufadel and Schlicker (2011) the term ***"wavelet"*** comes from a French term ***"ondelette"*** which means a ***"small wave"***. One can equally say that wavelets literally mean small waves, because they have finite length and are oscillatory. Lai (2015). The localized nature of wavelets enables them to be used in the analysis of episodic variations in the frequency composition or decomposition of data signals and they are thus referred to as a ***"mathematical microscope"*** Lai (2015). According to Boggess and Narcowich (2015), the desirable characteristic wavelets is that they can be used to ***"zoom in"*** on high frequency oscillations and still ***"zoom out"*** to model the underlying trend.

Among other criteria one might have to consider when choosing the wavelet basis functions, Hannonen (2006) and Boggess and Narcowich (2015) cite the following: (1) Symmetry (2) Orthogonality (3) Smoothness (4) Compact Support (5) Continuity

The Haar Wavelets

As an example, consider the definition of the Haar wavelet below:

The Haar Scaling Function

Definition 1 *The Haar scaling function is defined as*

$$\phi(x) = \begin{cases} 1, & if\ 0 \leq x < 1 \\ 0, & elsewhere \end{cases} \tag{1}$$

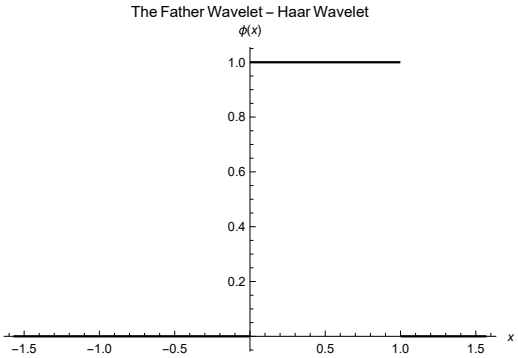


Figure 1: Haar Function

The Haar Wavelet

Definition 2 *The Haar wavelet is the function*

$$\psi(x) = \phi(2x) - \phi(2x-1)$$

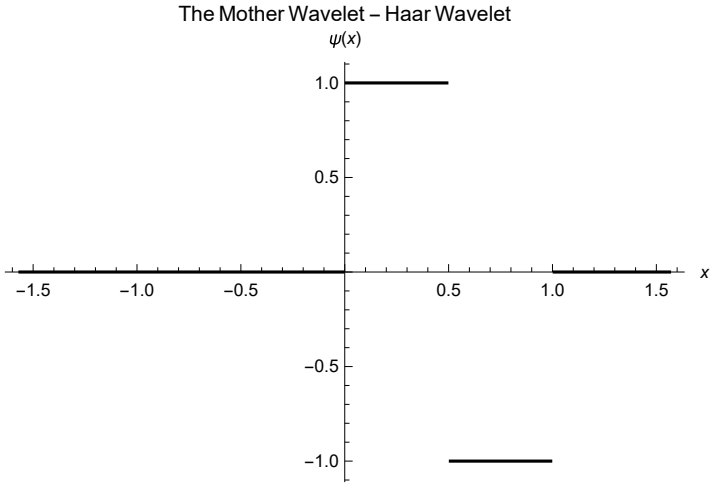


Figure 2: Haar Wavelet Function

Other Widely Used Wavelets

The Daubechies Wavelets

The Bettle-Lemarie Scaling Function

$$\phi(x) = \begin{cases} \frac{1}{2}x^2, & 0 \leq x < 1 \\ -x^2 + 3x - \frac{3}{2}, & 1 \leq x < 2 \\ \frac{1}{2}(x-3)^2, & 2 \leq x < 3 \\ 0, & otherwise. \end{cases}$$

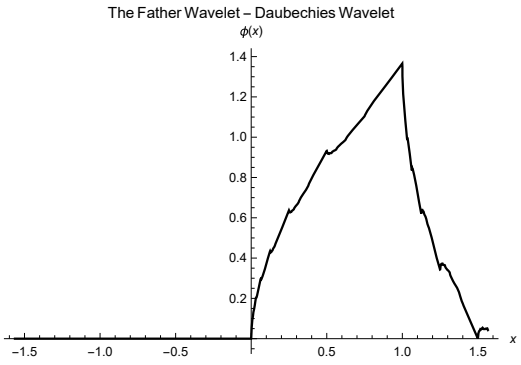


Figure 3: Daubechies Scaling Function

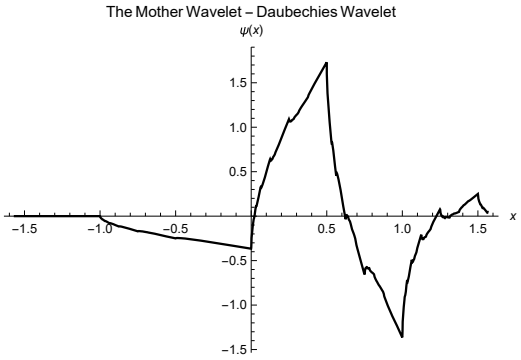


Figure 4: Daubechies Wavelet Function

The Shannon Scaling Function

$$\phi(x) = \begin{cases} \frac{sin(\pi x)}{\pi x}, & t \neq 0 \\ 1, & t = 0. \end{cases}$$

The Mexican Hat

Definition 3 The Mexican Hat Mother Wavelet *is defined by:*

$$\psi(x) = (1 - x^2)e^{(-\frac{t^2}{2})}$$

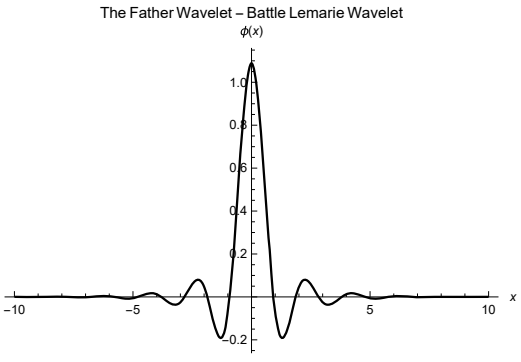


Figure 5: The Bettle-Lemarie Scaling Function

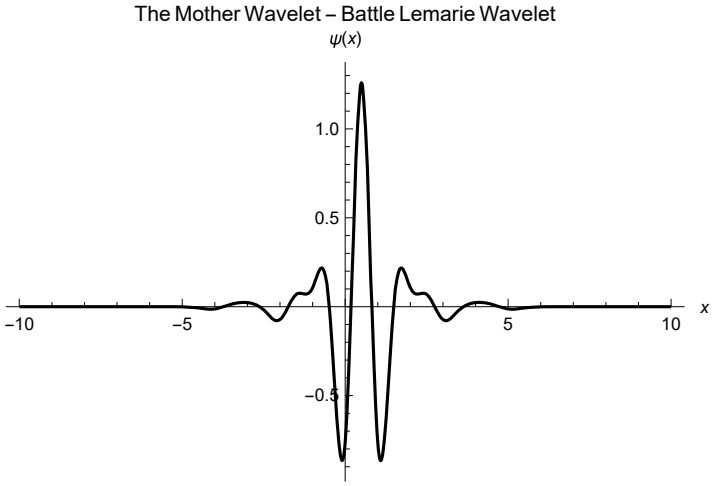


Figure 6: The Bettle-Lemarie Wavelet Function

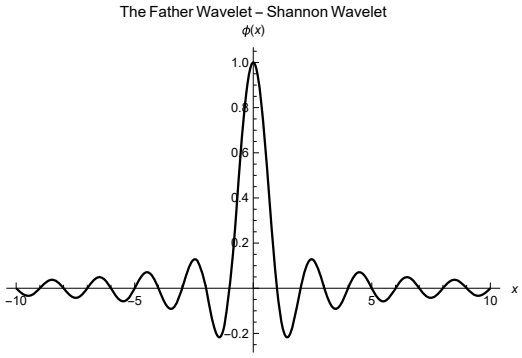


Figure 7: The Shannon Scaling Function

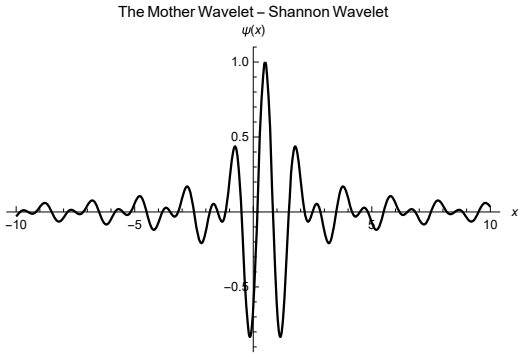


Figure 8: The Shannon Wavelet Function ®

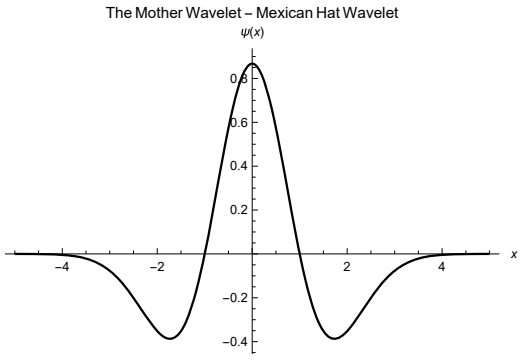


Figure 9: Mexican Hat Wavelet

The Paul Hat

Definition 4 The Paul Wavelet *is defined by:*

$$\psi(x) = \frac{(2i)^n n!(1-ix)^{-n-1}}{\sqrt{\pi}\sqrt{(2n)!}}$$

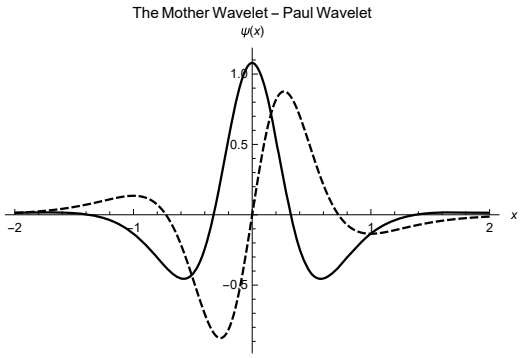


Figure 10: Paul Wavelet

METHODOLOGY

Quarter Circular Wavelet Basis

We develop a scalable wavelet basis by joining two quarter circles of varying radii. For theoretical details of the model, we assume that the two radii are equal, which we will relax in our curve fitting. Consider Figure 11 which shows how the idea of the quarter circular wavelet is developed from circle geometry.

The flexibility with this approach attributable to the idea that the two circles can have varying radii and the interval van be wide or narrow. In other words, all the basic properties of wavelet basis are indeed preserved. If $A = (x_1, y_1)$ and $B = (x_2, y_2)$, then the first approximation for $C = (\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2})$. The quarter-circular wavelet is ***infinitely differentiable*** at the points where the waves are joined. This is the overwhelming characteristic of this method, over the hierarchical, optimal dating method which produces lines when joining the peaks to the troughs. In other words, the wavelet in Figure 12 is perfectly horizontal at point A and perfectly vertical at point C.

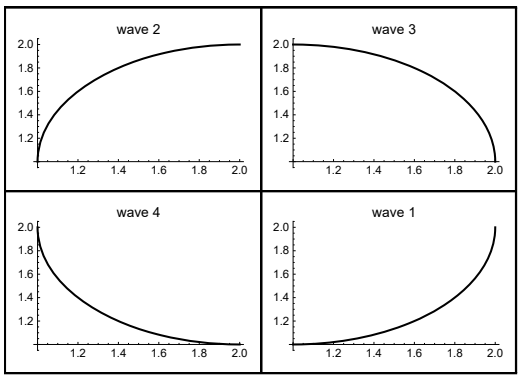


Figure 11: Quarter Circular Waves

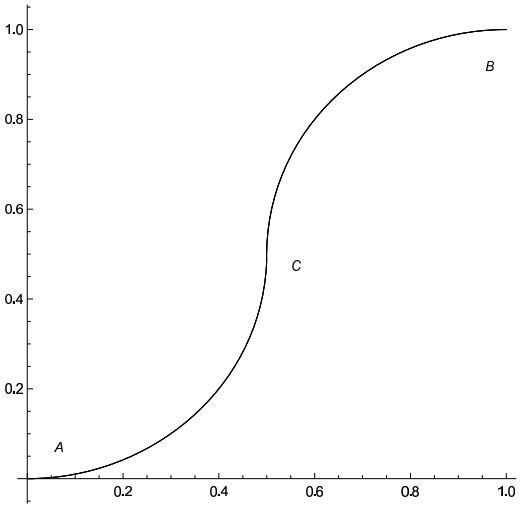


Figure 12: Quarter Circular Wavelets

Quarter Circular Wavelets

As illustrated in Table 1, the wave 1 begins where wave 4 ends, wave 2 begins where wave 1 ends, wave 3 begins where wave 2 ends and wave 4 begins where wave 3 ends. Effectively, one can notice that we only need to estimate two parameters $\{r_1, r_2\}$ to fit wave 1 and wave 2 and/or wave 3 and wave 4.

Table 1: Beginning and Ending of Quarter Circular Waves

wave 1	**wave 2**	**wave 3**	**wave 4**
$(x_c; y_c - r_2)$	$(x_c + r_1; y_c)$	$(x_c; y_c + r_2)$	$(x_c + r_1; y_c)$
$(x_c; y_c - r_2)$	$(x_c; y_c + r_2)$	$(x_c + r_1; y_c)$	$(x_c; y_c - r_2)$

Figure 13 shows the quarter circular wavelet, after concatenating wave 1, wave 2, wave 3 and wave 4 from Figure 11, in that order. Equivalently, combining wave 1 and wave 2 yield the wavelet in Figure 14, whereas wave 3 and wave 4 produce the wavelet in Figure 15. When the wavelets in Figure 14 and Figure 15 are concatenated, we have the ***"full cycle"***, as illustrated in Figure 13. Effectively, ***"trough to trough"*** or ***"peak to peak"*** defines a full cycle. To begin with wave 1, we chose the trough to trough definition of a full cycle.

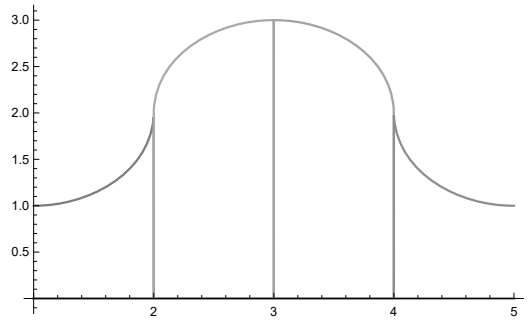


Figure 13: Quarter-Circular Wavelet **(Full Cycle)**

Although the waves in Figure 14 and Figure 15 appears symmetrical, this is only done for the simplicity of explaining the methodology. In principle, one can choose any radius for each of the waves and the waves will still join with ***"perfect smoothness"***. It will be seen that (for the practical examples shown) some of the fitted waves have very big radii while the other ones would have very small radii. Particulary, very high radius followed by relatively smaller radius would make the fitted curve to look kinked. Zooming the fitted model in Mathematica does, however, validate the fact that that the fitted model will always be smooth. Hence, the quarter-circular wavelet is very desirable, for cycle dating purposes.

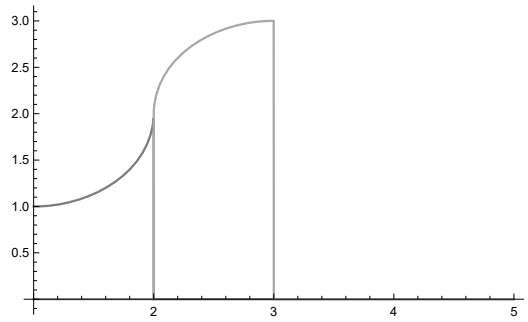


Figure 14: Wave 1 and Wave 2 **(Upward Wave)**

MODEL SPECIFICATION

We perform Non-Linear Model Fitting using ***"Wolfram Mathematica ®"***. The structure or form of the model is specified as a quarter-circle, accordingly. For convergence, we set some initial (starting) values for the end points of the various quarter-circles. This is described in the next section. We employ a method called the ***"Hierarchical Method"*** to divide the data range into smaller intervals where model fit would yiled local maxima and minima, based on optimizing some objective function, which typically is profit, for financial cycles. Unlike the quarter-circular approach,

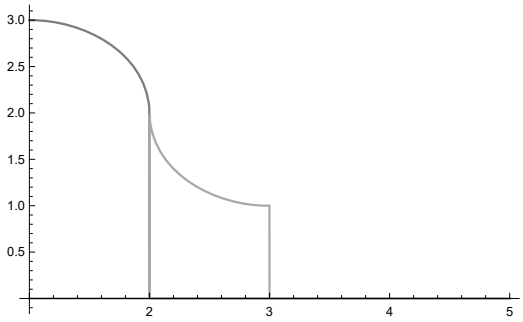


Figure 15: Wave 3 and Wave 4 **(Downward Wave)**

the Hierarchical method yields straight lines joining the peaks and troughs.

The similarity of the two two methods is that the number of peaks and troughs (which in essence measures the degree of smoothing or blurring) is pre-determined. Importantly, the quarter circles join smoothly, thus yielding a differentiable algebraic function upon which calculus can be employed to ascertain other special characteristics such as the turning points and the nature thereof. This, primarily, is the overarching advantage of the quarter-circular wavelet compared to the Hierarchical Method. It is needful that the values of A and B, through the data points, be found in a systematic, scientific manner. We employ the following methodology, as developed by Litvine (2016):

Initial Approximation: Hierarchical Method

Suppose we have a time-series of realizations $y(i), \quad i = 1, ..., n$ represents certain asset price traded at a stock exchange. Suppose we have an ideal trader who always purchases and sells the asset at the best possible times. Let $i_1, ..., i_m$ be the times at which the trades occur (i.e. $i_1, i_3, i_5, ..., i_{m-1}$ are times of buys and $i_2, i_4, i_6, ..., i_m$ be the times of sales and m is even). The profit that is achieved under the given $i_1, ..., i_m$, may be represented as:

$$OF = (y(i_2) + y(i_4) + y(i_6) + ... + y(i_m))$$
$$-(y(i_1) + y(i_3) + y(i_5) + ... + y(i_{m-1}))$$
$$\sum_{j=1}^{m/2}(y(i_{2j}) - y(i_{2j-1}))$$

This OF equals the total value of all sales less the total value of all buys. The ideal trader will maximize this function (OF) over all possible $i_1, ..., i_m$ and m:

$$\underset{i_1,...,i_m,m}{max} \quad OF$$

The above represents the maximum profit that potentially may be derived from trading the given asset over the defined period of time (the transaction costs are ignored at this stage).

Suppose we need to find cycles in this data in such a way that the given objective function (OF) is maximized. The method consists of the following steps.

1. Find such i_1 and i_2, that:
   - $1 < i_1 < i_2 < n$;
   - $OF(y(1), y(i_1), y(i_2), y(n))$ is maximized over all possible i_1 and i_2;
2. The points i_1 and i_2 divide the range (1, n) into 3 subintervals (we shall call them areas): $(1, i_1), (i_1, i_2)$ and (i_2, n);
3. Of the available areas select one which has the highest potential to have cycles (the examples of such selection techniques will be discussed later). We shall call this area a pivotal area.
4. The first point of this area will now be considered as point one, and the last point is the point number k;
5. Find such i_1 and i_2, that:
   - $1 < i_1 < i_2 < k$;
   - $OF(y(1), y(i_1), y(i_2), y(k))$ is maximized over all possible i_1 and i_2.
6. Now the pivotal area is divided into 3 subareas (by points i_1 and i_2 found at the present step) and the total number of areas in the series increase by 2.
7. Check the stopping criteria. If the stopping criteria is true, stop the algorithm. Repeat from step 3, while the stopping criteria is false.

NB: Every iteration of the above algorithm, the number of areas (and consequently the number of dividing points) will increase by two.

Steps for Fitting the Quarter Circular Wavelet

Now that we have explained how to find change points using the Hierachical method, we now turn the focus of this section to the step-by-step procedure for fitting the quarter-circular wavelets. Importantly, consider Figures 14 and 15. For every cycle, the endpoints of the respective upward wave or downward wave is evaluated using the Hierachical method, as expalined in the section above. Therefore, the Quarter circular method seeks to find two parameters (x,y), the midpoint of each wave (upward or downward). This indeed enhances computational efficiency. Particularly, we use the built-in Nonlinear model fit algorithm of Wolfram Mathematica ®.

Economic Cycles

In this section, we consider two examples to illustrate how the proposed methodology works for the dating of economic cycles. Section presents the quarter circular non-linear model fit for the Consumer Price Index of Energy in South Africa. The second example on economic cycles is based on the Civilian Unemployment Rate in the United States of America. Section presents the model fit, accordingly. For every cycle, we present the original data series, the optimal dating model and the quarter circular non-linear model fit. For a pictorial presentation of the goodness of fit, we also provide a transposition of the quarter-circular non-linear model over the time series. Formal statistical measures of goodness of fit will be presented in Section .

Quarter Circular Non-Linear Model Fit: Consumer Price Index - Energy - South Africa

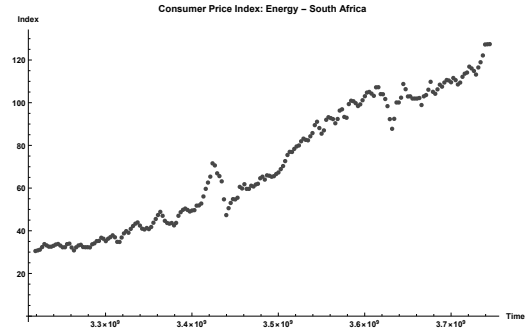


Figure 16: Consumer Price Index of Energy in South Africa - **Original Series**

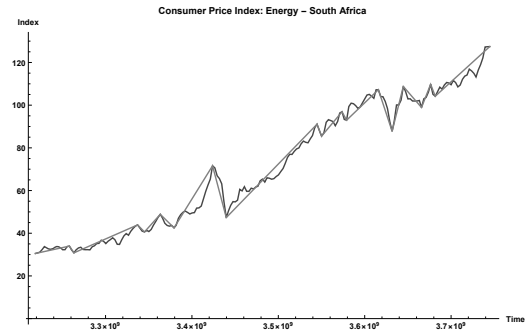


Figure 17: Consumer Price Index of Energy in South Africa - **Optimal Dating**

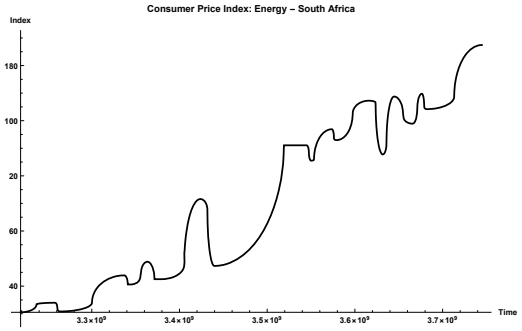


Figure 18: Consumer Price Index of Energy in South Africa - **Quarter Circular Model Fit**

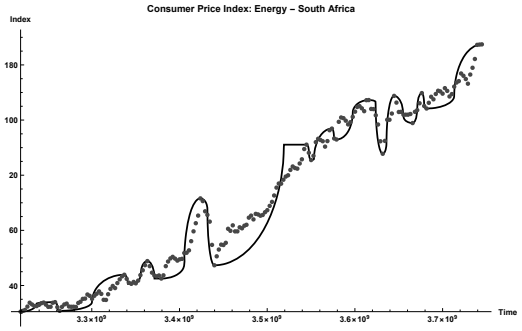


Figure 19: Consumer Price Index of Energy in South Africa - **Quarter Circular Model Fit vs. Data**

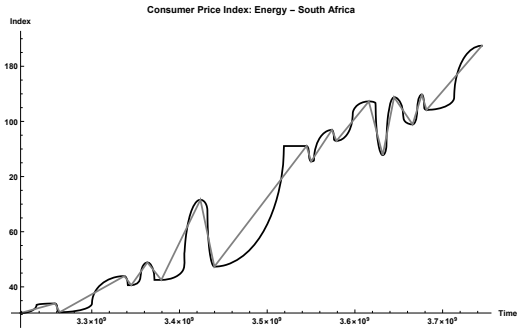


Figure 20: Consumer Price Index of Energy in South Africa - **Quarter Circular Model Fit vs. Optimal Dating**

Quarter Circular Non-Linear Model Fit: Civilian Unemployment Rate - United States of America

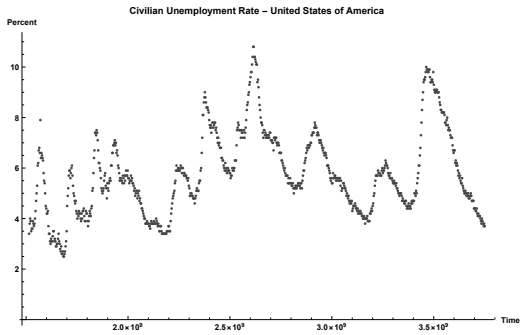


Figure 21: Civilian Unemployment Rate - United States of America - **Original Series**

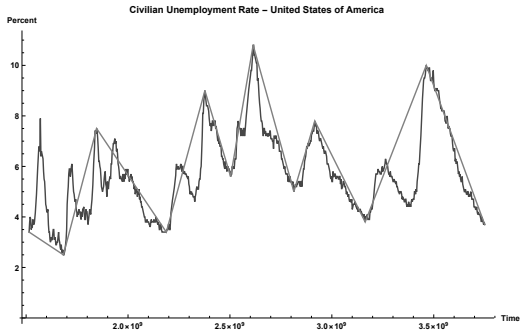


Figure 22: Civilian Unemployment Rate - United States of America - **Optimal Dating**

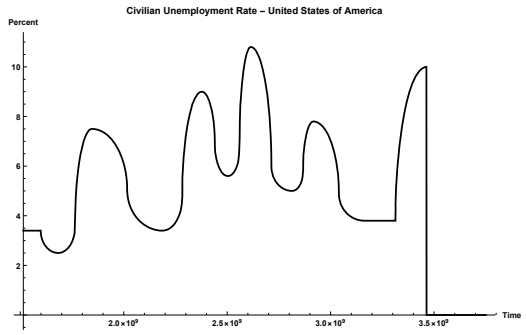


Figure 23: Civilian Unemployment Rate - United States of America - **Quarter Circular Model Fit**

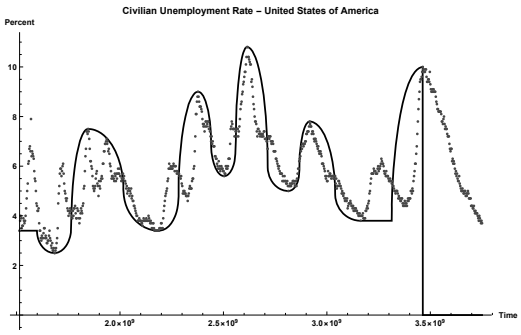


Figure 24: Civilian Unemployment Rate - United States of America - **Quarter Circular Model Fit vs. Data**

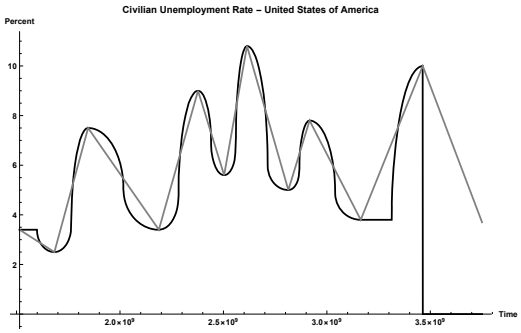


Figure 25: Civilian Unemployment Rate - United States of America - **Quarter Circular Model Fit vs. Optimal Dating**

Financial Cycles

We consider two examples to illustrate how the proposed methodology works for the dating of financial cycles. Section presents the quarter circular non-linear model fit for the exchange rate of the South African rand to the United States dollar. The second example on financial cycles is based on the South African interest rates on government bonds. Section presents the model fit, accordingly. For every cycle, we present the original data series, the optimal dating model and the quarter circular non-linear model fit. For a pictorial presentation of the goodness of fit, we also provide a transposition of the quarter-circular non-linear model over the time series. Formal statistical measures of goodness of fit will be presented in Section .

Quarter Circular Non-Linear Model Fit: South Africa - United States - Exchange Rate

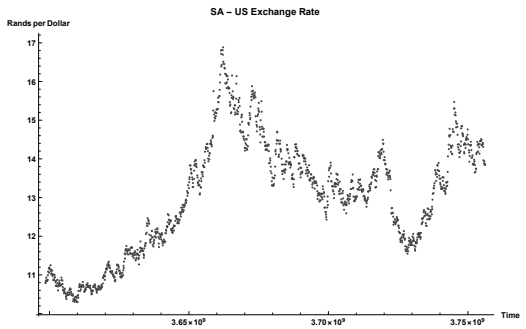


Figure 26: South Africa - United States - Exchange Rate - **Original Series**

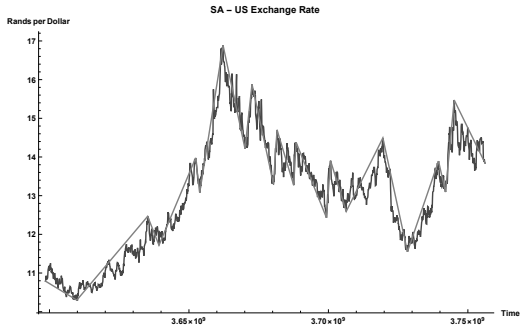


Figure 27: South Africa - United States - Exchange Rate - **Optimal Dating**

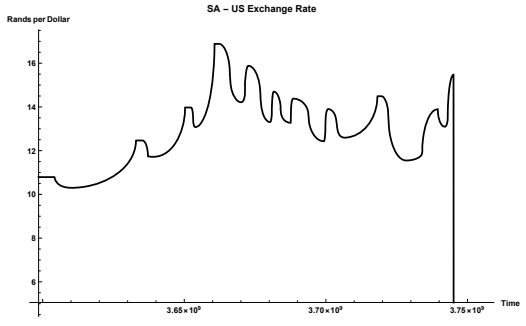


Figure 28: South Africa - United States - Exchange Rate - **Quarter Circular Model Fit**

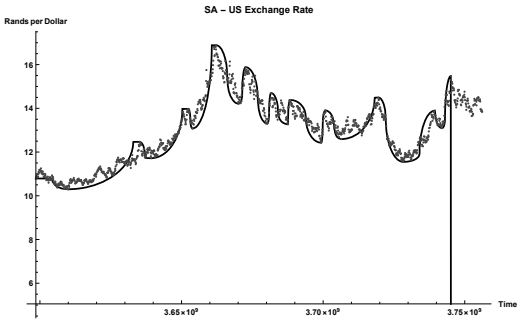


Figure 29: South Africa - United States - Exchange Rate - **Quarter Circular Model Fit vs. Data**

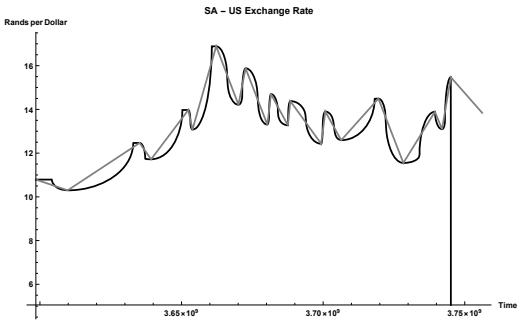


Figure 30: South Africa - United States - Exchange Rate - **Quarter Circular Model Fit vs. Optimal Dating**

Quarter Circular Non-Linear Model Fit: Interest Rates Government Bonds - South Africa

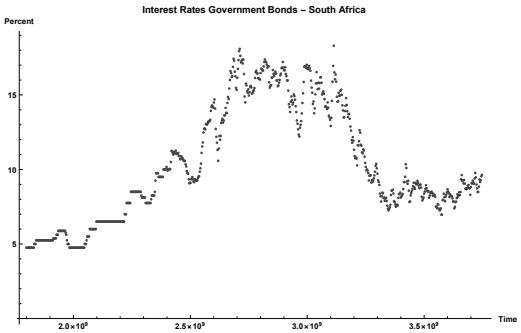


Figure 31: Interest Rates Government Bonds - South Africa - **Original Series**

Goodness of Fit: Quarter Circular Non-Linear Model

In this section, we present a summary of the goodness of fit statistics for the various models. Particularly, we employ the adjusted R^2, a measure of the percentage of variability in the data "explained" by the quarter-circular waves.

Economic Cycles

We present the goodness of fit statistics for the quarter circular non-linear models of the economic cycles. The

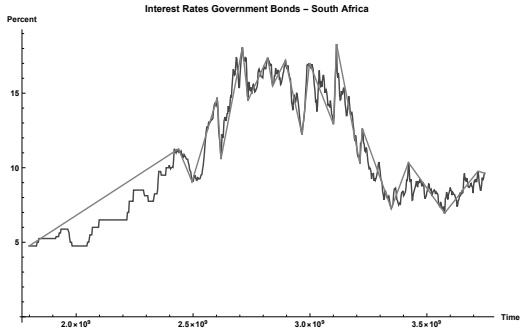


Figure 32: Interest Rates Government Bonds - South Africa - **Optimal Dating**

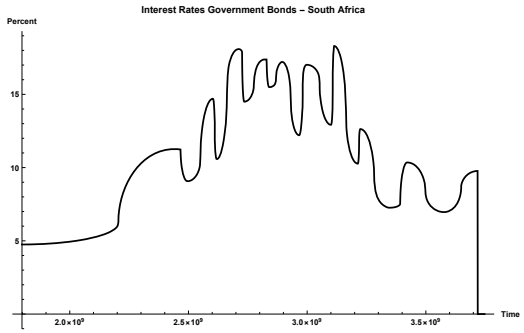


Figure 33: Interest Rates Government Bonds - South Africa - **Quarter Circular Model Fit**

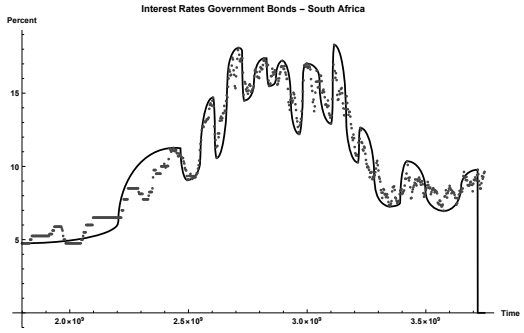


Figure 34: Interest Rates Government Bonds - South Africa - **Quarter Circular Model Fit vs. Data**

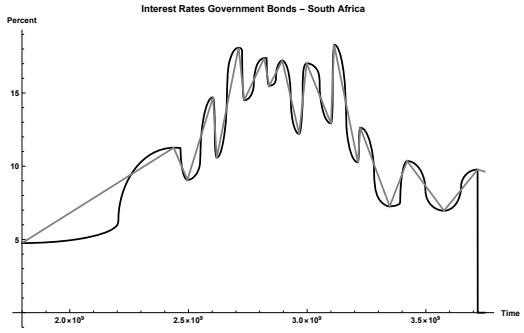


Figure 35: Interest Rates Government Bonds - South Africa - **Quarter Circular Model Fit vs. Optimal Dating**

goodness of fit for the ***"Upward Waves"*** waves and the ***"Downward Waves"*** are presented in Tables 3, 4 and Tables 2, 5, respectively.

Consumer Price Index: Energy - South Africa (5 cycles)

Table 2: Consumer Price Index: Energy - South Africa - Adjusted R^2 for the fitted **"Downward Waves"**

Model	1	2	3	4	5
Adj R^2	0.92	0.94	0.79	0.87	0.94

Table 3: Consumer Price Index: Energy - South Africa - Adjusted R^2 for the Fitted **"Upward Waves"**

Model	1	2	3	4	5
Adj R^2	-	0.09	0.53	0.82	0.88

Civilian Unemployment Rate - United States of America (5 cycles)

Table 4: Civilian Unemployment Rate - United States of America - Adjusted R^2 for the Fitted **"Upward Waves"**

Model	1	2	3	4	5
Adj R^2	0.76	0.80	0.93	0.96	0.75

Table 5: Civilian Unemployment Rate - United States of America - Adjusted R^2 for the Fitted **"Downward Waves"**

Model	1	2	3	4	5
Adj R^2	0.87	0.96	0.98	0.95	0.97

Financial Cycles

We present the goodness of fit statistics for the quarter circular non-linear models of the financial cycles. The goodness of fit for the ***"Upward Waves"*** and the ***"Downward Waves"*** are presented in Tables 6, 9 and Tables 7, 8, respectively.

South Africa - United States - Exchange Rate (5 cycles)

Table 6: South Africa - United States - Exchange Rate - Adjusted R^2 for the Fitted **"Upward Waves"**

Model	1	2	3	4	5
Adj R^2	0.99	0.99	0.98	0.94	0.89

Table 7: South Africa - United States - Exchange Rate - Adjusted R^2 for the Fitted **"Downward Waves"**

Model	1	2	3	4	5
Adj R^2	0.99	0.97	0.92	0.98	0.99

Interest Rates Government Bonds - South Africa (5 cycles)

Table 8: Interest Rates Government Bonds - South Africa - Adjusted R^2 for the Fitted **"Downward Waves"**

Model	1	2	3	4	5
Adj R^2	0.96	0.96	0.95	0.96	0.94

Table 9: Interest Rates Government Bonds - South Africa - Adjusted R^2 for the Fitted **"Upward Waves"**

Model	2	4	6	8	10
Adj R^2 0.96	0.90	0.97	0.96	0.97	

Summary

In this study, a scalable, smooth and differentiable quarter-circular wavelet basis for the smoothing of business, economic and financial time series is proposed. A new method for the dating of cycles is proposed, by bringing wavelets, circle geometry and non-linear regression analysis together. The models are implemented in Wolfram Mathematica ®, which affords efficient algorithms for symbolic, numerical and functional computation of the proposed methodology. Practical illustrations are given, from South African and other international indices. The development of wavelet bases from circle geometry was pivotal to this study. Circular waves indeed can be used for the modeling of cycles. From the wheel, which in no doubt is one of the greatest innovations of human-kind, this study resembles another further development from the concept of a circle.

Indeed this work could be seen as the first suggestion of quarter-circular wavelets, bringing together theory from spectral analysis and geometry. The novelty of the proposed methodology might not be in question. What could be more interesting is the fact that the word cycle possibly have the same root as the word circle. This could require a linguistic approach through the history of languages. This study was in a way methodological. Various statistical, mathematical and econometric concepts and constructs were articulated in this work. These include - Fourier analysis, waves, wavelets, the Hurst Exponent, filters, causality and fractals. Importantly, these were seen as tools and aspects of methodology applicable for the explication of business, economic and financial cycles. Most of the data used to in the study pertains to the South African economy. Nevertheless, considering that the data was utilized entirely for illustrative purposes, the usage of indices from other economies could as well be materialize.

Future Work

Applications of the Quarter-Circular Wavelet Model

The method proposed in this work could safely be applied to the study of cycles in other non-financial domains. Examples of such applications could include geophysical cycles and medical cycles. The advantage with the approach developed herein is that one can always choose new bases functions, depending on the research problem. This is reserved and/or recommended as future work. Particularly, with the surge of statistical research methodologies in medical science, dating of peaks and troughs in biometric data could indeed resemble an advancement in quantitative medical research. Furthermore, the proposed methodology could see an effective application in seismic signal processing of data in mining explorations. This is said, considering that bases such as the Morlet wavelet were actually introduced for the analysis of seismic signals.

Choice of the Number of Cycles

In this work, we intuitively select the number of cycles to consider. For practical purposes, one might need to have a scientific way of choosing the number cycles. An example could be to use a ratio of the the theoretical duration of the particular cycle to the interval of enquiry. For example if it's known that a particular cycle lasts at least 3 years e.g. the Kitchin cycle in Table 10 and one is analyzing data spanning over 60 years, one could reasonably expect at most $\frac{60}{5} = 20$ cycles.

Forecasting Peaks and Troughs

Given the dating of cycles, one can utilize the dating information in building forecasting models for economic and financial cycles. This could be of significant inter-

Table 10: Economic Waves

Cycle/wave name	Period (years)
Kitchin cycle	3 - 5
Juglar cycle	7 - 11
Kuznets swing	15 - 25
Kondratiev wave	45 - 60

est to investors, policy makers and fund managers. For example, establishing that a particular cycle generally occurs after say 5 years would mean that the next peak or trough would be expected to occur in 2020 if the last observed peak or trough was in 2015. Nevertheless, the distribution of the duration of cycles could be stochastic.

Cross-Wavelet Cycle Synchronization

Considering that the current cross-wavelet analysis employed bi-variate spectral analysis in economics and finance is based on the Morlet wavelet (which was originally developed for geophysical signals), developing the method proposed in this paper further for cycle synchronization would in doubt resemble scientific progress in the statistical modeling of economic and financial cycles.

REFERENCES

Aboufadel E. and Schlicker S., 2011. *Discovering wavelets*. John Wiley & Sons.

Boggess A. and Narcowich F.J., 2015. *A first course in wavelets with Fourier analysis*. John Wiley & Sons.

Hannonen M., 2006. *An analysis of trends and cycles of land prices using wavelet transforms. International Journal of Strategic Property Management*, 10, no. 1, 1–21.

Lai W., 2015. *Wavelet theory and its applications in economics and finance*. Ph.D. thesis, Department of Economics.

Litvine N.I., 2016. *Economic and Financial Cycles in South Africa*. Ph.D. thesis, Université de Lorraine.

OPTIMAL DATING OF CYCLES

Alpheus Mahoya and Igor Litvine
COEF, Nelson Mandela University
Port Elizabeth 6031, South Africa
email: mahoyashingi@gmail.com

KEYWORDS
Forex, Financial Cycles, Large Data, Hierarchical, Forecasting

ABSTRACT

Financial investors often face an urgent need to predict the future and this calls for accurate forecasting. Financial time series forecasting is an interesting research issue that has received growing attention in recent years. Reliable forecasting is therefore a fundamental aspect in the decision making field. However, financial markets are characterised by various shocks and this makes forecasting very complex. Litvine and Francis (2015) proposed an approach for dating cycles in financial time series for the purposes of optimising buy/sell strategies. In this approach, cycle dating is presented as an optimisation problem. Another method by Litvine and Francis (2015) for optimising this problem, known as the hierarchical method (using full evaluation 2, or HR-FE2) was also introduced. This method is partially impractical for large data sets because of it's longer computational time. In this paper we employ this hierarchical method to derive our peaks and troughs for our cycles. A wave (data between two adjacent buy/sell points) will be selected and a differentiable sinusoidal wave with certain parameters will be used to fit this selected wave. This will be repeated for all waves and a full sequence derived. A statistical comparison was then made between these methods, including HR-FE2. Analysis of Goodness of Fit, R^2, AIC and also the computational time on different time series lengths, Hurst exponent, and number of buy-sell points.

INTRODUCTION

The foreign currency exchange (forex) market is a decentralised, over-the-counter and interbank world financial market where participants buy, sell, exchange and speculate on currencies Garner (2012). The forex market is the biggest financial market with daily transactions exceeding five trillion dollars Ozturk et al. (2016). The forecasting of forex prices is of great importance to many investors, speculators, practitioners and scientists but the chaotic and volatile structure of the price in the forex market complicates forecast analysis. There are two main approaches in analysing forex prices namely, the fundamental analysis approach which uses macro-economic factors to explain and forecast price changes and the technical analysis approach which uses historical data to forecast future prices. The technical analysis approach can be further divided into chart analysis and technical indicator analysis. Chart analysis seeks to find recurring patterns in charts. Technical indicator analysis seeks to measure the volatility in price, detect patterns, identify trends and define the correlation between price and volume by transforming historical time series prices into another time series data Ozturk et al. (2016). This study will follow the technical analysis approach.

Cycles are characterised by accidental (random) rhythm fluctuations which are similar to regularly alternating up and down forces Dewey and Mandino (1971). These ups and downs (peaks and troughs) can make it difficult to find meaning in the cycles, especially when dealing with large data sets. For example, in time series data (with reference to the economics) forex prices can fluctuate several times in a minute, rendering the use of cycles quite tedious (especially for a full trading day). To counter this, Litvine and Bismans suggested that identifying overall downward and upward trends between segments in the the cycles would be more meaningful. This process involves maximising the objective function to find alternating peaks and troughs that separate the upward and downward trend segments and it is known as *cycle dating* and is illustrated in Figure 1. The troughs and peaks are referred to as buy-sell points respectively Litvine and Francis (2015). The optimisation problem is presented as follows:

Maximise
over $x_i^*, i = 1, 2, ..., K$ (where K is even)

$$f_{profit}(X^*) = \sum_{i=1}^{K/2} (y(x_{2i}^*) - y(x_{2i-1}^*))$$

subject to

$$x_i^* < x_j^*, i < j,$$

where K is the number of buy/sell points selected, $y(x)$ is the forex price at time index x and X^* is the set of time indices of buy/sell points, i.e. $X^* = \{x_1^*, x_2^*, ..., x_K^*\}$.
The dating of cycles is important for the following reasons Litvine (2016) :

1. Data reduction. Investors are not really concerned about the whole cycle but certain attributes of the cycle e.g. the height/magnitude of the cycle (measured in currency) and the length of the cycle (measured in time).

2. Data smoothing. This is a noise reduction process that allows only the important patterns of the cycle to stand out. This is important because outliers are thus minimised from the pattern to avoid bias.

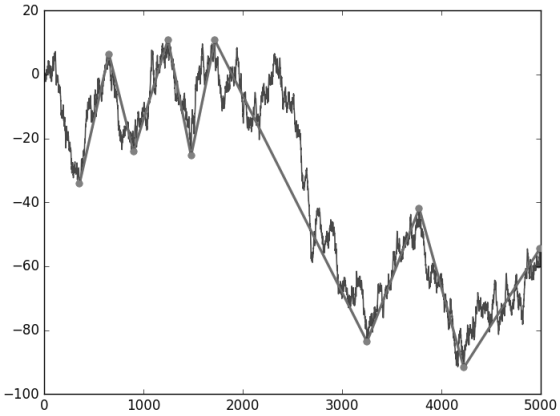


Figure 1: An example of dated cycles where "buy" and "sell" points extracted from the time series are the shown dots

3. Benchmarking and comparing trader performance. From past cycles, trained algorithms can devise optimal strategies that allow for maximum possible profits. This potential profit compared to the actual profit the trader gets can give an overall assessment of the trader's performance.

4. Time-series data is often measured on different time scales, still one may need to simultaneously study two series recorded at different intervals. Classical analysis (e.g. standard correlation analysis) requires that both sets be on the same scale and have equal length. This requirement is unnecessary for comparing cycles (e.g. for synchronisation analysis).

Different cycle dating techniques perform differently on different cycles Boldin (1994), it is important therefore to distinguish which of these techniques is the best for reaching the optimal solution. This is because the time taken to reach this optimal solution is crucial as time allows investors/-traders to act swiftly in certain instances. Thus, if more computational time is required to reach optimal solutions then bureaucratic decision making processes take longer, resulting in lost revenue/potential profit. So the overall aims of this study are firstly to determine how different cycle dating methods reach optimal solution. Secondly to find out which of these techniques is the most effective and thirdly to attempt to further improve the method that was deduced to be the best one. This can be done incorporating certain characteristics of the other alternative methods analysed.

PREVIOUS STUDIES

The first detailed analysis of self-similar cycles was conducted by Ralph Nelson Elliott in the 1930s Elliott (1994). The Elliott Wave Principle refers to the theory that social behaviour drifts back and forth in recognisable patterns Prechter (1994). Elliot recognised, using stock market data, that five-wave and three-wave movement patterns occur repeatedly in market price data. He further described how these waves are combined together to form larger versions of the same pattern, how they then link together to form identical patterns of the next large size and so on.

Frost and Prechter explained the rules and process of identifying these patterns suggested by Elliot Frost and Prechter (1995). They suggested that five-wave patterns are impulsive waves and follow the trend of the larger degree wave while three-wave patterns are corrective waves and go against the larger degree trend. They also mention several variations in the basic impulsive and corrective wave pattern. Based on these patterns, they can forecast the most probable direction and magnitude of the next wave to occur, thus also predicting the peaks and troughs that arise at the end of those waves. Figure 2 illustrates those concepts.

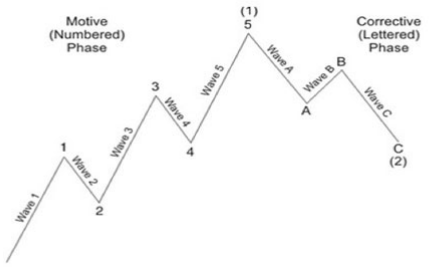


Figure 2: Elliott's 5-3 Cycle: The longer wave is known as the motif and the shorter wave known as the correction.

After the Elliot waves in the 1930s, Burns and Mitchell Burns and Mitchell (1947) then defined cycles as waves that inevitably evolve from one into another: prosperity, crisis, depression and revival. They suggested that in any time series, a cycle is found by identifying points of rises (peaks) and falls (troughs). This process is known as "cycle dating".

Bry and Boschan Bry and Boschan (1971) took the work of Burns and Mitchell Burns and Mitchell (1947) further and introduced a computational algorithm that isolates the turning points by identifying the peaks and troughs of the analysed series. The algorithm is well known as the BB-algorithm and is illustrated by Figure 3.

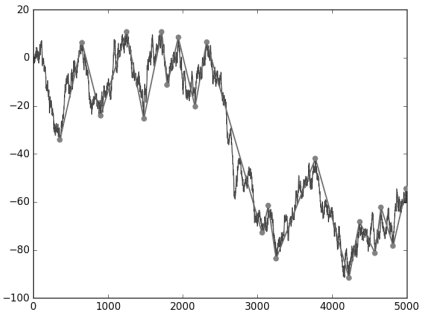


Figure 3: Example of dating cycles in simulated data with BB algorithm.

The BB-algorithm was more suitable for economic data with a sampling period of only a month or so (Kapp 2017) but the drawback is the fact that most economic data is only available

quarterly, semi-annually and annually. This work was then taken further by Pagan Pagan et al. (2013) who made the algorithm suitable to quarterly data and this was termed the BBQ-algorithm. The drawbacks of both the BB and BBQ algorithms is that they both only derive the length of time the growth/contractions occur but ignore the magnitude of this change. This is a huge disadvantage to traders who aim to buy and sell forex with maximum profit over a short period. A modification was then done to the BBQ algorithm by James Engle Pagan et al. (2013) so as to make it a more reliable and versatile algorithm and this was known as the MBBQ-algorithm.

Stock market data has a very high sampling frequency, for example, a time series of forex prices (exchange rates) typically has a sampling interval of less than a second. This poses a very big problem to the application of BB, BBQ and MBBQ algorithms to such data because they were specifically designed for economic data, and for monthly and quarterly intervals between observations. This is because the BB-type algorithms have a criteria of using horizontal windows of neighbouring observations in identifying cycles, rather than the vertical distance between observations (Kapp 2017).

The above-mentioned drawbacks of the BB-type algorithms drove the merging of statistical tools and computer science concepts into developing algorithms that can deal with high sampling frequency and volatile data (Kapp 2017). Wu and Huang Wu and Huang (2009) developed an algorithm that holds no assumptions on the sampling frequency. This algorithm uses a similar technique to the BB-type algorithms and represents a time series with the most prominent peaks and troughs in the series. These prominent peaks and troughs are found in a sequential manner and this allows for stopping conditions when the number of peaks and trough extracted are enough. This is something a BB algorithm cannot do.

In the modern algorithms, the dating procedure is performed by optimising an objective function whose domain consists of discrete time series observations and is not continuous in nature. Many optimisation methods rely on some continuity and differentiability to work (Kapp 2017) and therefore they cannot be used. Stochastic optimisation algorithms are methods that either optimise objective functions that contain random noise, or that search for optimal solutions using random variables Spall (2004). These methods do not require any continuity or differentiability in the objective function. Among the stochastic optimisation algorithms are Genetic algorithms (GAs), differential evolution algorithms (DE) and partial swarm optimisation algorithms (PSO).

Litvine and Bismans Litvine and Francis (2015) proposed a class of methods, known as the hierarchical methods, that were designed to find cycles in data by optimising an objective function. This study will aim to introduce new algorithms that might be considered alternatives to the hierarchical methodology, or that improve on the existing hierarchical methods. The improvements desired are firstly higher values obtained for the specified objective function, and secondly, lower computation times.

DATA AND METHODOLOGY

The data used for this study and on which the algorithms were applied to is the USD/ZAR currency pair prices. The data obtained were daily prices of this pair from January 1980 to April 2019 totalling to 10239 observations. Only data from March 2014 to February 2019, a total of 1306 observations, was used for sampling purposes. A highly computational software package called WOLFRAM MATHEMATICA (version 12) will be employed to run all the analysis.

Waves

One of the main objectives on any times series is to reduce noise and smooth the time series. In this section we introduce continuous differentiable functions which optimally remove noise and smooth it thus revealing cycles. According to Litvine (2016) a basic wave is defined as a function $y(x)$ which is defined on the interval $0 <_x < 1$, and satisfies the following:

1. is continuous
2. is differentiable
3. is increasing
4. $y(0) = 0$;
5. $y(1) = 1$;
6. $y^{'}(0) = 0$;
7. $y^{'}(1) = 0$;

In this paper, a basic Sin Wave was employed for fitting purposes. This choice was comprehended by the fact that a basic Sin Wave meets the above-mentioned criteria and also it's power of flexibility. A basic Sin Wave is defined by the following:

$$y(x) = \frac{1}{2} + \frac{1}{2}sin(\pi(x - \frac{1}{2}) = \frac{1}{2} + \frac{1}{2}sin(\pi x - \frac{\pi}{2})$$

or alternatively

$$y(x) = \frac{1}{2} - \frac{1}{2}cos(\pi x) = sin^2(\frac{nx}{2})$$

and is illustrated in Figure 4.

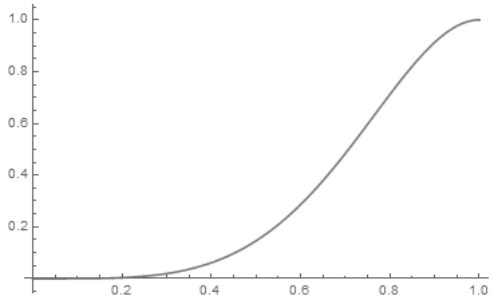


Figure 4: Basic Sin Wave

Fitting

This section illustrates how the Sin Wave will be applied to smooth out the original time series. For this purpose, the entire series will be divided into section as follows: the first section will be the series between the first corresponding buy and sell points, the second section being the series between the last section's end point (buy or sell) and the corresponding buy or sell point. This is done for entire series.

To show how the Sin Wave will be used to fit these sections of the entire series, a wave was simulated and is show below:

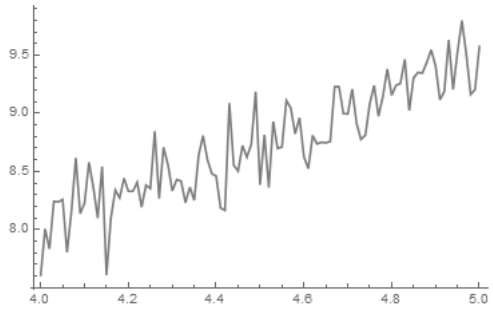


Figure 5: Simulated Wave

The sinusoidal wave defined was then fitted to this random wave to remove the noise but still maintaining the two end points and the following was obtained.

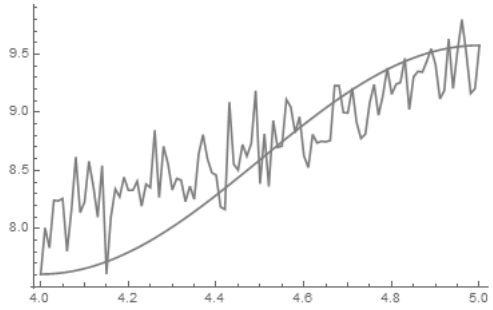


Figure 6: Sin Wave fitted on a time series

It is evident that the wave was able to extract the impurities and irregularities in the series but still maintaining the long term trend of the series.

RESULTS

The Sin Wave was fitted to all sections of the series and the waves were then joined together to form a smoothed series. Below in Figure 7 is a graphical illustration of the original series against the smoothed wave.
The wave managed to remove the noise as evident in the Figure above. However, from the period of 1980 to 2001 there wasn't much noise in the series which means forex prices were not really too volatile although there was an upward long-term trend. Much noise was from 2001 to 2019 we witness high volatility in the forex prices and still the wave managed to smooth the series. To get more insight, we zoomed into the series from wave 25 to 35 and is illustrated in Figure 8 below.

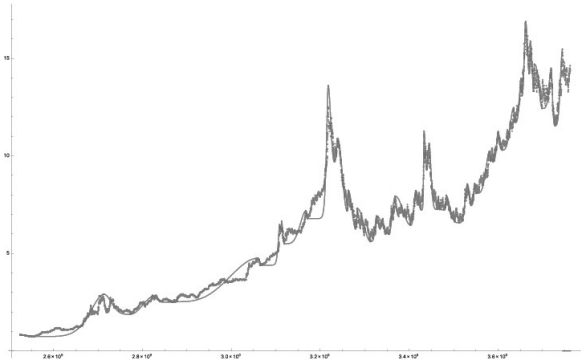


Figure 7: Original Series plotted against the Smoothed Wave

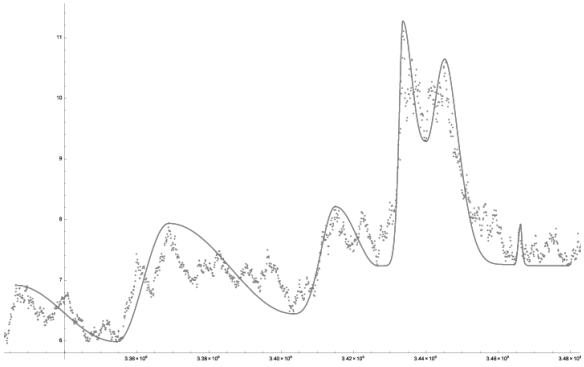


Figure 8: Zoomed image of between wave 25 to wave 25

From this picture, it is more clearer how the noise in the original series was effectively removed. Since volatility is very crucial in time series as it also represents risk, the derivative of the fitted wave was deduced and is graphically illustrated in Figure 9 below.

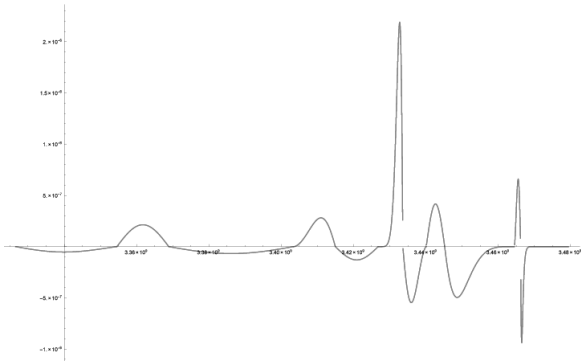


Figure 9: Derivative of the Fitted Wave

The higher the derivative of the wave from the horizontal axis, the higher the volatility in the prices at that particular period. It is evident from the figure above to deduce where prices were highly volatile and these periods were 1984-1987, 2000-2003, 2007-2010 and also 2016-2019. The highest volatility was evident at the period 2000-2003.

CONCLUSION

The fitted Sin Wave was able to smooth the original series and removed most of the noise without loosing the important data. As indicated, a basic Sin Wave should be differentiable.

This showed to be a very important and unique aspect of this wave of choice, more insight into the original series was archived and periods of high and low volatility identified. This is important insight as it gives investors a reflection of the risk associated which periods of high volatility. The next and important objective is to be able to parametrise this fitted wave, find the distribution of these parameters and use this to predict the next possible buy/sell point. This is a possible research are in this regard.

REFERENCES

Boldin M.D., 1994. *Dating turning points in the business cycle*. *Journal of Business*, 97–131.

Bry G. and Boschan C., 1971. *Programmed selection of cyclical turning points*. In *Cyclical analysis of time series: Selected procedures and computer programs*, NBER. 7–63.

Burns A.F. and Mitchell W.C., 1947. *Measuring Business Cycles*. *Science and Society*, 11, no. 2, 192–195.

Dewey E.R. and Dakin E.F., 1949. *Cycles, the science of prediction: with 1950 postscript: Further deflation and its promise*. Holt.

Dewey E.R. and Mandino O., 1971. *Cycles: The mysterious forces that trigger events*. Hawthorn Books.

Elliott R.N., 1994. *RN Elliott's Masterworks: The Definitive Collection*. New Classics Library.

Fidrmuc J. and Korhonen I., 2010. *The impact of the global financial crisis on business cycles in Asian emerging economies*. *Journal of Asian Economics*, 21, no. 3, 293–303.

Frost A.J. and Prechter R.R., 1995. *Elliott wave principle: key to market behavior*. New Classics Library.

Garner C., 2012. *Currency trading in the forex and futures markets*. FT Press.

Hamilton J.D., 1989. *A new approach to the economic analysis of nonstationary time series and the business cycle*. *Econometrica: Journal of the Econometric Society*, 357–384.

Harding D. and Pagan A., 2002. *Dissecting the cycle: a methodological investigation*. *Journal of monetary economics*, 49, no. 2, 365–381.

Kapp K.P., 2017. *Optimal Cycle Dating of Large Financial Time Series*. Master's thesis, Nelson Mandela Metropolitan University.

Kele T.; Lombaard C.; Mouton S.; and van der Merwe L., 2010. *Elementary Statistics for Business and Economics*. Heinemann. South Africa.

Litvine I. and Francis B., 2015. *Dating financial cycles with hierachical method*. 57th Annual South Africa Statistical Association Conference.

Litvine I.N., 2016. *Economic and Financial Cycles in South Africa*. Ph.D. thesis, Université de Lorraine.

Ozturk M.; Toroslu I.H.; and Fidan G., 2016. *Heuristic based trading system on Forex data using technical indicator rules*. *Applied Soft Computing*, 43, 170–186.

Pagan A. et al., 2013. *Patterns and their uses*. *National Centre for Econometric Research Working Paper*, , no. 96.

Prechter R., 1994. *RN Elliott's Masterworks: The Definitive Collection*. *New Classics*.

Spall J., 2004. *Stochastic Optimization, Handbook of Computational Statistics*.

Tsay R.S., 2005. *Analysis of financial time series*. John Wiley & Sons.

Wu X. and Huang D., 2009. *Representing financial time series based on important extrema points*. In *2009 Third International Symposium on Intelligent Information Technology Application*. IEEE, vol. 1, 501–504.

AUTHOR BIOGRAPHY

ALPHEUS MAHOYA is a Zimbabwean born student at Nelson Mandela University in South Africa where he commenced his studies in 2014. Having completed his bachelor degree in Economics and Statistics in 2016, he also went on to complete his Honours degree in Mathematical Statistics in 2017. Alpheus is currently a second year Masters student and an Associate Lecturer at Nelson Mandela University.

MARKETS TREND ASSESSMENT USING NATURAL LANGUAGE PROCESSING

Justin Mudzimu and Igor Litvine
COEF, Nelson Mandela University
Port Elizabeth 6031
South Africa
E-mail: justinmudzimu@gmail.com

KEYWORDS
natural language processing, financial cycles, large data, sentiment analysis, machine learning

ABSTRACT

The study of natural language processing has been a lot active in recent decades, however it was only in recent years that more statistical approaches have been developed and been used in the field of finance. This study is follows an approach for analysing co-occurrence matrices. The data that is used in this study was scrapped from the web to make a corpus from finance news articles text. The scrapped text was used to make the co-occurrence matrices. The approach this study uses relies on co-occurrence matrices to produce heat-map images that can be visualised and interpreted using machine learning methods. Methods used in this study are practical for large data sets and as in the case of this study, these data sets may be acquired over a lengthy period of time in cases of primary data collection. This study will rely on machine learning techniques to classify image outputs from co-occurrence matrices computed from our corpus.

INTRODUCTION

The key problem faced in finance, be it on an individual or institutional level is that of knowing with an acceptable degree of confidence when and where to invest. This problem results in a lot of effort and funding being put in, in-order to try to beat the market mostly through efficient use of available information. A question to be asked would be how much information is available and how well are we able to use this information for informed decision making. The turn of the century brought about a digital era in which constant improvement in digital technologies resulted in exponentially increasing amounts of data which, however messy, has presented our generation with an opportunity to improve our livelihood. There are a lot of news websites on the internet today with each of these news site publishing hundreds and more articles daily.

Most ordinary individuals, financial analysts and advisers look out for news articles to try to detect trends in the market and by so doing attempt to profitably decide when to buy, hold or sell securities. The advancement in the digital world came with an increase in machine computational power. Methods have been developed to handle large text data that we may extract from the web and pass through a machine that in turn reads the hoards of text, make sense of them and aid with interpretation that is free of human error and bias. This field of work has been termed natural language processing(NLP), and it is an artificial intelligence field that lies at the intersection of computing science, linguistics and statistical methods.

PREVIOUS STUDIES

Natural Language Processing is not a very old technique and one could argue that it was born in the 1950s when one mathematician, Alan Turing, asked the question, "Can Machines Think?". This question in his paper, *Computing Machinery and Intelligence*, is what we can say marked the pioneering of modern day artificial intelligence. The test widely known today as the Turing test, tests for machine intelligence. The test suggests that an interviewer asks questions to concealed interviewees, a machine and a human being. The machine would be said to have passed the test if the interviewer can not tell that they just spoke to a machine, Alan (1950).

In his 1950 paper Turing had already anticipated the difficulties of developing A.I by manually programming a digital computer. His suggested remedy was that machines must learn in the same way as a human child, Muggleton (2014). Turing was at this point anticipating or better put giving birth to what is modern day machine learning. Little was to be done in his time as Turing cited computational hardships, a position that was quite justifiable.

While Turing made significant contributions to the future of A.I, major contributions were being done in the field of linguistics by John R Firth and Zellig Harris. Firth wrote, "...the complete meaning of a word is always contextual, and no study of meaning apart from a complete context can be taken seriously", Firth (1935). In 1954, at the forefront of research in linguistics, Harris wrote a paper titled the "distributional structure", the ideas of which would be seen being used today in computational linguistics.

Harris suggested that there exists some distributional structure in natural language, spoken or written. Elements in any language do not occur arbitrarily relative to each other, Harris (1954). Our utterances or writings are not put together by our free choices of words to put across whatever message we intend to mean but in actual fact, words have to be put in some order, an order that regularly occurs, for our words to bear any meaning at all,Harris (1954). As if he was anticipating deep learning, Socher et al. (2013), Harris

went on to suggest that the occurrences of language elements are a network of interrelated statements, some of which are a result of a derivation from other statements and not from distributional order restriction of individual elements, Harris (1954). Harris' work was echoed by J R Firth who received a lot of recognition as he built on top of Harris' distributional hypothesis which is now known by most through Firth's words: "You shall know a word by the company it keeps" Firth (1957).

The distributional approach utilizes distributional properties of linguistic entities as the building blocks of semantics. The general approach behind the hypothesis is quite clear, it suggests that there is a strong correlation between distributional similarity and semantic similarity which allows us to use the former to estimate the latter Sahlgren (2008).

However great, early approaches to the study of the distributional hypothesis-co-occurrence analysis-rested upon the manual identification of collocations in relatively small amounts of text which were limited in scope and coverage, Bartsch and Evert (2014).

A more statistical approach to the distributional hypothesis coupled with the increase in computational power in the early 1980s brought about a huge improvement in natural language processing. It is suggested that statistical patterns of human word usage can be used to understand what words mean, Furnas et al. (1983). Christopher Potts clarifys the suggestion in the paper by Furnans and authors by saying that if text units appear to have similar vectors in a text frequency matrix, it is an indication of similarity in meaning (semantic similarity) of the text units, Socher et al. (2013).

Text frequency matrices are as a result of different word embedding techniques. The most common and simplest would be the bag of words technique(BoW). According to Clark (2014), in this bag of words approach, the frequency of words is accounted for and nothing else. An example of a word frequency matrix is given below.

Table 1: Term-term co-occurrence matrix

Term	*w1*	*w2*	*w3*	*w4*	*w5*	*w6*
w1	0	1	0	0	0	0
w2	0	0	1	0	0	3
w3	1	0	0	2	0	0
w4	3	0	0	1	1	0
w5	0	1	3	0	1	2
w6	1	2	0	0	0	0

Some words that have little information may appear more than the ones more relevant. This sort of problem may be addressed by dividing term frequencies by their document frequency to give a term frequency-inverse document frequency(tf-idf). The tf-idf helps to reweigh term frequencies such that those that appear in about every document and hence less discriminatory have the least weights, Clark (2014).

In a different approach, Deerwester suggested the application of singular value decomposition to the term-document matrix, a method they called latent semantic analysis/latent semantic indexing (LSA/LSI). The method decomposes the large term-document matrix into an smaller estimate matrix of the large term-document matrix. LSA is an improvement to the bag of words approach as it allows terms to be paired with documents that do not necessarily contain the exact words contained in the term. This allows for different words in the terms to be paired with a document made of corresponding synonymous words. This method brings similar documents closer together in the vector space while separating more different document, Deerwester et al. (1990). Other methods can be used to the same effect as the SVD for semantic analysis.

Another approach, co-occurrence analysis vectors, appreciates the early works of Harris and Firth on the distributional hypothesis and considers a word's neighbouring words. In this approach, a specific window size is set. Each word will be defined by its neighbouring words to the left as well as the one to the right. This is modelled mathematically by constructing a co-occurrence matrix for each window. In this study, a sentence window is considered as it provides a more logical unit so as not to miss essential co-occurrences due to the arbitrarily choosing a window.

The use of the co-occurrence analysis technique, gives a co-occurrence matrix, the result of which can be visualised as a heat-map that can be rearranged to distinguish chaotic matrices from information bearing ones, Ryabchenko et al. (2015). These heat-maps can be used through machine learning image classification to distinguish between different context messages in articles and in the case of this study, positive information(bullish information) from negative information(bearish).

DATA AND METHODOLOGY

In conducting this study, the data to which the classifying algorithms were applied to was collected for the first time by scraping a South African financial news website, Sharenet, using a scraper code written in Mathematica. The data collected was raw, so it had to go through some cleaning before any training was done. All the numerical values in the articles were removed, extra space and punctuation was removed. All the words were separated, and sentences were converted into lists and co-occurrences were calculated over a sentence window, over a single day's publications. The choice of the website was because the site is rich with local news and would be ideal for making an assessment of the local markets. Sharenet seemed to have a consistent web layout so it was reasonable to scrap data from Sharenet as there would be hardly any need to rewrite the scraper.

Although most authors favour the removal of stop words, no stop word removal was done in the cleaning of the collected data. Even though stop words are said to be non information bearing, they are necessary for grammar, give meaning to

publications and do contribute to the structure of an article, in terms of the distributional hypothesis. After all this pre-processing, the text was ready for vectorisation.
A word embedding (vectorising) technique of choice was used, there are several of those available, including the bag of words approach, TF-IDF and latent semantic analysis amongst others, Velay and Daniel (2018). The corpus that was made from the pre-processed data was used to make a $100 * 100$ matrix of vector values from a sentence window term-term co-occurrences. Co-occurrence

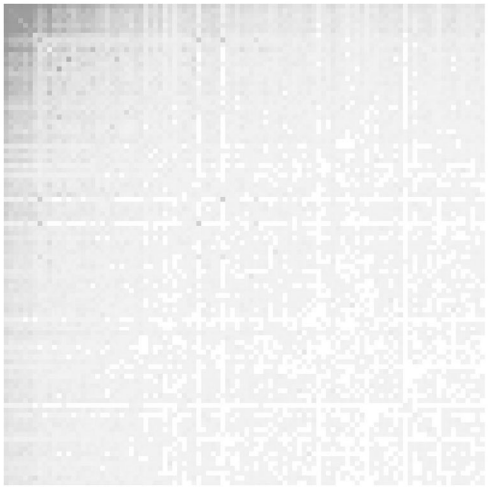


Figure 1: "A 100×100 fragment co-occurrence matrix heat-map"

analysis is favoured in conducting this study as it naturally produces superior matrix heat-maps that can be used for image classification, Ryabchenko et al. (2015). This thesis study is conducting a context similarity analysis based on co-occurrence matrices.

This study favours heat-maps as they present us with visuals that about everyone can be presented with and given a brief explanation to what information they bear and they would easily understand, unlike being presented with pages of text or math. The heat map representation of our news articles context will allow us to easily associate a specific context with an expansionary or recessionary market.

Mathematica software was used for all the analysis and programming codes used in this study. Mathematica has a total of 10 integrated machine learning algorithms. This study will take advantage of some of those algorithms to classify the heat-maps mentioned before. There is a general consensus in the A.I field that there is no universal superior machine learning algorithm. Given that no superior classifier exists, there is enough reason to use a number of the learning classifiers and then choose the best out of how they perform with the out of sample data.

Mathematica has integrated learning algorithms that include NearestNeighbors, Logistic Regression, Naive Bayes, Decision Tree, Random Forest and Neural Networks.

Support Vector Machines, Decision Trees (Random Forest) and Regression models have dominated classification models in the past decades with over 70 percent of previous studies having used these models. Regression models are particularly preferred since we can explicitly observe the impact of each factor included and analyse the importance of variables by dropping them out, Xing et al. (2018). SVMs have their roots in statistical learning theory, Cortes and Vapnik (1995) and have gotten so much popularity owing to their being robust, accurate and producing good results even when exposed to little training data, Anthony et al. (2007). Support vector machine are a supervised learning technique that analyses data and identify the pattern used for classification. Given the training data, SVMs build a maximum hyperlane that separates the data in to two categories. This makes the SVMs binary linear models in nature, Jain (2013). However, one drawback of linear models is that they rely on strong hypotheses, for example, a Gaussian distribution of independent variables, which is not always the case in real world cases, hence other algorithms like Neural networks and the Naive Bayes are always used, Xing et al. (2018).

The Naive Bayes is a probabilistic classifier that takes off from the statistical Bayes Theorem. Like SVMs the Naive Bayes classifier can also produce efficient results with small training data. In addition to being accurate the Naive Bayes is also computationally fast compared to other learning methods. One good trait this algorithm possess is that of a short learning time as it is free of iterative procedures. The Naive Bayes is built on a strong assumption of independence between features. The independence assumption does not hold sometimes in the real world but the classifier however yields optimal results even under circumstances when the independence assumption does not hold, Park (2016). Machine learning algorithms have different pros and cons and in this study more than four will be used and the results will be compared to get an optimal classifier.

RESULTS

This study has as of now used a total of over 1200 articles with over 300 000 words. More data is being collected by the day to get a wider sample space and hence reach more conclusive arguments. The results discussed here are over 20 days of news paper articles.

Using Mathematica, 20 100×100 co-occurrence matrices were computed and from them, 20 image representations (heat-maps) of news data were created. The resultant images were applied for training on some of the machine learning algorithms, results of which are given below. It so happened that of the 20 days data being analysed here, the South African markets were evenly split between first 10 days where markets were generally going down and a remainder of 10 days when markets were generally stable, slightly going up. This here instead of thousands of words gave us a perfect set of a boolean data set, markets going down/otherwise, that could easily be classified. Classification accuracy over this training data is given in the results table below. Given the amount of data used, the results are encouraging as more data will only mean improved results.

Table 2: Learning Methods Accuracy

Methods	Accuracy
Logistic Regression	0.995
NearestNeighbors	1
RandomForest	0.990
DecisionTree	0.895
NeuralNetwork	0.555
Naive Bayes	1

These results are quite positive as they prove that machine learning algorithms can distinguish between different co-occurrence matrices. This result is what was anticipated and with more data being gathered, investigations on how rearranging these matrices can help with visualising the differences between images from these different sets of data.

CONCLUSION

This study attempts to give investors a competitive edge through automating the search of vital free existing textual information (mostly in the form of news articles) using natural language processing. An individual can not attempt to asses and pick up vital trading information from gigabytes of text data that is available to them on a daily basis. Fortunately we have machines with high computational specifications lately and through natural language processing methods, we can have a competitive edge over fellow investors using the superior information retrieval methods available to us. This study attempts to use a combination of machine learning text processing and image classification techniques to reach the desired goal. From presented findings, it is recommended to use Naive Bayes and NearestNeibhors as the preferred learning methods, with the possibility of using Logistic Regression as it also gives a good accuracy result. It would be interesting, given enough data to create enough heat maps to have a visual representation of different trend days for comprehension purposes to those not familiar and not interested in the finer details of the study involved but are just interested in the results. Further studies would be to attempt to use the results of this study create an NLP based indicator that reflects the current state of the assessed financial markets. This indicator, if successful can possibly be used to forecast market trends.

REFERENCES

Alan T., 1950. *Computing machinery and intelligence-AM Turing*. *Mind*, 59, no. 236, 433.

Anthony G.; Gregg H.; and Tshilidzi M., 2007. *Image classification using SVMs: one-against-one vs one-against-all*. *arXiv preprint arXiv:07112914*.

Bartsch S. and Evert S., 2014. *Towards a Firthian notion of collocation*. *OPAL - Online publizierte Arbeiten zur Linguistik*, 2014, 48–61.

Clark S., 2014. *Vector Space Models of Lexical Meaning*. Wiley Blackwell.

Cortes C. and Vapnik V., 1995. *Support-vector networks*. *Machine learning*, 20, no. 3, 273–297.

Deerwester S.; Dumais S.T.; Furnas G.W.; Landauer T.K.; and Harshman R., 1990. *Indexing by latent semantic analysis*. *Journal of the American society for information science*, 41, no. 6, 391–407.

Firth J.R., 1935. *The Technique of Semantics. Transactions of the philological society*, 34, no. 1, 36–73.

Firth J.R., 1957. *A synopsis of linguistic theory, 1930-1955*. *Studies in linguistic analysis*.

Furnas G.W.; Landauer T.K.; Gomez L.M.; and Dumais S.T., 1983. *Human factors and behavioral science: Statistical semantics: Analysis of the potential performance of key-word information systems*. *The Bell System Technical Journal*, 62, no. 6, 1753–1806.

Harris Z.S., 1954. *Distributional structure*. *Word*, 10, no. 2-3, 146–162.

Jain S., 2013. *A machine learning approach: Svm for image classification in cbir*. *Internationa Journal of Aplication or Annovation in Engineering & Management (IJAIEM)*, 2, no. 4.

Muggleton S., 2014. *Alan Turing and the development of Artificial Intelligence*. *AI communications*, 27, no. 1, 3–10.

Park D.C., 2016. *Image classification using Naïve Bayes classifier*. *Int J Comp Sci Elec Eng*, 4, no. 3, 135–139.

Ryabchenko O.; Litvine I.; and Dibrova A., 2015. *Biosocial Economy as a Mechanism for Transition to Sustainable Development*. *Demography and social economy*, 161–172. doi:10.15407/dse2015.03.011.

Sahlgren M., 2008. *The distributional hypothesis*. *Italian Journal of Disability Studies*, 20, 33–53.

Socher R.; Perelygin A.; Wu J.; Chuang J.; Manning C.D.; Ng A.; and Potts C., 2013. *Recursive deep models for semantic compositionality over a sentiment treebank*. In *Proceedings of the 2013 conference on empirical methods in natural language processing*. 1631–1642.

Velay M. and Daniel F., 2018. *Using NLP on news headlines to predict index trends*. *CoRR*, abs/1806.09533. URL `http://arxiv.org/abs/1806.09533`.

Xing F.Z.; Cambria E.; and Welsch R.E., 2018. *Natural language based financial forecasting: a survey*. *Artificial Intelligence Review*, 50, no. 1, 49–73.

SMART GRID SIMULATION

AN OPTIMIZED LINEAR REGRESSION MACHINE FOR PREDICTION AND INTEGRATION OF RENEWABLE SOURCES IN SMART GRID.

Leila Ziouche
Computer Science Departement,
Faculty of Mathematical, Physical and Natural Sciences,
University of Tunis-El Manar, Tunisia.
Email: leila.ziouche@fst.utm.tn

Syrine Ben Meskina
Computer Science Departement,
ESPRIT, School of Business,
University of Carthage, Tunisia.
Email: Syrine.benmeskina@gmail.com

Mohamed Khalgui
National Institute of Applied Sciences and Technology,
University of Carthage, Tunisia.
Email: khalgui.mohamed@gmail.com

Laid Kahloul
LINFI Lab, Computer Science Department,
Biskra University, Algeria.
Email: laid.k.b@gmail.com

KEYWORDS
Smart grid, Failures, Recovery Sources, Linear Programming, Simplex, Linear Regression, FROSLR.

ABSTRACT

To enhance the recovery of enormous number of failures of electric grids and guarantee best power availability during technical teams interventions, we have proposed a strategy that predicts the number of renewable sources needed for faults recovery by respecting total required loads and given budget by ministry. In this paper, we improve the accuracy, feasibility and efficiency of reported solution based on Linear Programming and Linear Regression. Thus, we have implemented a tool for Renewable Sources Prediction for Failures Recovery called -FROSLR-, we test and validate obtained results. We conclude by investigating the hypothesis stating that the increasing of the renewable energy guides to get best power availability and minimize clients waiting time.

I. INTRODUCTION

Recent researches on smart grids have integrated renewable and green energies to save big buildings, hospitals and research laboratories. They employed management services to enhance recovery and maintenance (ZAFAR et al. 2017). Multiple failures of power electric networks were able to engender another problems that need fast recovery by emergency sources, besides the intervention of technical teams. Abidi et al. (2017) proposed an approach to balance between production and consumption of energy and make the verification of power availability using smart agents. While, Nefedov et al. (2018) encouraged the use of electrical vehicles to cover big buildings, they complemented a mathematical and simulation model to estimate saved energy.
Moreover, Ben Meskina et al. (2018) implemented a tool for faults detection using smart agents, isolation of malfunctioned lines to recover failures using emergency lines. While, Ben Meskina et al. (2017) developed a reconfiguring methodology of smart grids. Hong et al. (2010) applied multiple linear regression for load forecasting in the United States. Also, they discussed the power demand based on temperature and temperature square per hour. Ziouche et al. (2019) have proposed a cloud architecture of smart grids rebuilding for technical teams intervention scheduling to fix a defined number of failures. The least number of intervened technical teams define a problem of delayed assisting to repair failures with highest priorities what may cause a miss of electricity for some regions that results economic losses.
Thereby, recorded failures in smart grid can be classified into recovered and non recovered failures according to the availability of emergency sources where during the maintenance of this kind of failures, there exist some problems that differ from an electric grid to another. The discussed problems are listed as follows:

- Inability to cover sufficiently the total required load by recorded critical failures during maintenance.
- Definition clients required load per hour, day, week and month without to be able to satisfy that need in case of blackouts, hard interruptions and other kinds of failures.
- The least number of technical teams that assist for smart grid failures repairing and late intervention that creates a miss of electricity for a long time.
- The increasing of total waiting time of critical failures during the intervention of technical teams.

These problems result huge losses in terms of time, amount of energy and finance. Also, these failures may engender other dominant and equivalent faults. So that, the deployment of new renewable energy sources has the ability to enlarge and enhance recovery. In this paper, we aim to control perfectly the sources of energy in a smart grid. So, we propose a prediction model based on Linear Programming and Linear Regression that forecasts the number of renewable sources required by using the historic data of smart grids what allows to:

- Give the highest priority of coverage and maintenance to critical failures.
- Define threshold number of renewable sources to be integrated in order to recover the maximum required load of critical failures.
- Enhance the power availability in case of interruptions and outages to avoid blackouts and losses.
- Reduce the total waiting time of clients without electricity during technical teams interventions.

We implement this heuristic in python to simplify the work of technical teams at real-time, validate and demonstrate the accuracy and effectiveness of this strategy in terms of power availability and total waiting time. The paper is organized as follows: Section II proposes a formalization of the requirements. Section III discusses the proposed contribution to resolve presented problem. Section IV considers an experimental study and improves obtained gains after implementation. Section V concludes with a discussion of future works.

II. REQUIREMENTS FORMALIZATION

In this section, we formalize the proposed smart grid model and prediction methodology separately.

A. Smart Grid Model

A smart grid SG is expected as a power network graph with different components and lines, i.e.,

$$SG = G(C, L) \tag{1}$$

where C and L are respectively the sets of electrical components and lines. C includes a set of transformers that distribute energy to a set of consumers, i.e.,

$$C = TRANS \cup CONS \tag{2}$$

where $TRANS$ is the set of medium and low transformers (MVT, LVT) and $CONS$ is the set of medium and low consumers (MC, LC) in SG. Also, L is composed of principal lines and emergency lines, i.e.,

$$L = PL \cup EL \tag{3}$$

where PL is the set of principal lines that are permanently activated and EL is the set of emergency lines that are activated only in case of recovery.

1. Recovery Sources

A smart grid SG includes a set of pertinent emergency lines coming from a central power plant PS and a global set of renewable sources RS that are used for failures recovery. RS represents all the solar panels, wind turbines and batteries sets, i.e.,

$$RS = PV \cup WTF \cup BT \tag{4}$$

where $PV = \{pv_1, ..., pv_{N_1}\}$ is the set of solar cells, $WTF = \{wtf_1, ..., wtf_{N_2}\}$ is the set of wind turbines and $BT = \{bt_1, ..., bt_{N_3}\}$ is the set of batteries, with N_1, N_2 and N_3 represents respectively the numbers of solar panels, wind turbines and batteries in SG.

The distributed renewable sources are designed as follows: (i) Every renewable source $pv \in PV$ or $wtf \in WTF$ is defined by its produced energy E_{pv} or E_{wtf}, and recovery time RT_{pv} or RT_{wtf}, (ii) Batteries $bt \in BT$ stores produced power by power plant or renewable sources to be used for failures recovery, it is defined by its storage capacity C_{bt}, available energy E_{bt} and recovery time RT_{bt}. The total recovery time $TRTA_{RS}$ and total available energy TRA_{RS} of existing renewable sources are computed as in Equation (5).

$$\begin{aligned} TRTA_{RS} &= RT_{pv} * N_1 + RT_{wtf} * N_2 + RT_{bt} * N_3 \\ TRA_{RS} &= E_{pv} * N_1 + E_{wtf} * N_2 + E_{bt} * N_3 \end{aligned} \tag{5}$$

2. Failures

A smart grid SG can suffer from a global set of failures $SFL(SG)$ that may encounter a component or line x_k, i.e.,

$$SFL(SG) = \{f(x_1), .., f(x_{Nf})\} \tag{6}$$

where Nf is the number of recorded failures in SG.

A failure $f(x_k) \in SFL(SG)$ is defined by a: (i) required load $ReqL_{f(x_k)}$, (ii) set of recovery sources $s(f(x_k))$, (iii) recovery time $RT_{s(f(x_k))}$, (iv) fixing time $FT_{f(x_k)}$, and (v) waiting time $WT_{f(x_k)}$. The total required load by a global set of failures is defined in Equation (7).

$$TReqL_{SFL(SG)} = \sum_{k=1}^{Nf} ReqL(f(x_k)) \tag{7}$$

However, $SFL(SG)$ includes two sets of failures: critical failures $CF(SG)$ or non critical failures $NCF(SG)$, i.e.,

$$\begin{aligned} &SFL(SG) = NCF(SG) \cup CF(SG) \\ &NCF(SG) = \{f(x_k) \in SFL(SG), E_{s(f(x_k))} \geq ReqL_{f(x_k)}\} \\ &CF(SG) = \{f(x_k) \in SFL(SG), E_{s(f(x_k))} < ReqL_{f(x_k)}\} \end{aligned} \tag{8}$$

(i) Non critical failures $NCF(SG)$: is the set of failures that have the least priority thanks to the power sufficiency and availability for recovery using available sources $E_{s(f(x_k))} \geq ReqL_{f(x_k)}$, (ii) Critical failures $CF(SG)$: refers to a set of failures that suffer from the unavailability or the insufficiency of energy where $E_{s(f(x_k))} < ReqL_{f(x_k)}$, they have the highest priority to be maintained and repaired first by technical teams.

3. Technical Teams

They are specialized experts that intervene for failures repairing in SG at run-time. Every smart grid includes a defined number of teams $N_{tt} \neq 0$. It allows to define the total waiting time $TWT_{SFL(SG)} = \sum_{k=1}^{Nf} \alpha_k * WT_{f(x_k)}$ and total fixing time $TFT_{SFL(SG)} = \sum_{k=1}^{Nf} FT_{f(x_k)}$ of

the set of failures inserted in TT_j. $FT_{f(x_k)}$and $WT_{f(x_k)}$ are the fixing time and waiting time of the k^{th} failure and α_k is the number of repetition of fault per month.

B. Prediction Machine Model

The investigation of the provided heuristic for problematic resolution includes three steps:

1. Linear Programming
Employs as input: the number of critical failures N_{CF}, total required load of critical failures $TReqL_{CF}$, total recovery time needed by critical failures TRT_{CF}, total fixing time TFT_{CF} and total waiting time TWT_{CF} during a month. It provides as output: the numbers of solar panels, wind turbines and batteries, i.e.,

$$\begin{cases} Max\ N_{RS} = (N_1 + y_1) + (N_2 + y_2) + (N_3 + y_3) \\ P_{pv} * y_1 + P_{wtf} * y_2 + P_{bt} * y_3 \leq Tcost \\ E_{pv} * y_1 + E_{wtf} * y_2 + E_{bt} * y_3 \leq a \\ RT_{pv} * y_1 + RT_{wtf} * y_2 + RT_{bt} * y_3 \leq b \\ y_1, y_2, y_3, N_1, N_2, N_3 \geq 0 \end{cases} \tag{9}$$

where N_{RS} is the number of renewable sources to maximize, y_1, y_2 and y_3 are respectively the numbers of new added solar panels, wind turbines and batteries. P_{pv}, P_{wtf} and P_{bt} refers to the prices in dollar of a solar panel, a wind turbine and a battery. E_{pv}, E_{wtf} are in order, the produced loads by a solar panel and a wind turbine, and E_{bt} is the capacity of battery, RT_{pv}, RT_{wtf} and RT_{bt} are respectively the recovery times of a solar panel, a wind turbine and a battery, and $a = TReqL_{CF}$ - TRA_{RS} and $b = TWT_{CF}$+TFT_{CF} -$TRTA_{RS}$ are the total needed load and the total recovery time of failures.

$$\begin{cases} (N_1 + y_1) + (N_2 + y_2) + (N_3 + y_3) = N_{RS} \\ P_{pv} * y_1 + P_{wtf} * y_2 + P_{bt} * y_3 + e_1 = Tcost \\ E_{pv} * y_1 + E_{wtf} * y_2 + E_{bt} * y_3 + e_2 = a \\ RT_{pv} * y_1 + RT_{wtf} * y_2 + RT_{bt} * y_3 + e_3 = b \\ y_1, y_2, y_3, N_1, N_2, N_3, e_1, e_2, e_3 \geq 0 \end{cases} \tag{10}$$

We use Simplex Method to obtain normalized constraints described in Equation (10) above. We have e_1, e_2 and e_3 are the slack variables and they permit to adjust given restrictions (Burkard et al. 2009).The first row of initial table of Simplex consists of the objective function coefficients, while the last row includes the objective function value and integrated number of renewable sources N_{NRS}=y_1+y_2+y_3 The rest values are decision coefficients and slack variables. The obtained N_{NRS_i} values are stored in a training set Y where $i = 1, .., n$ with n is the number of months $n \geq 3$.

2. Linear Regression
After getting the training set Y that refers to the number of renewable sources in a number of months, we apply Linear Regression Algorithm to predict the numbers of renewable sources for the rest of testing months.

(i) In order to choose the set of predictors X, we compute the best fit line based on total required load and the number of critical failures. We choose β_0 and β_1 that has the least error ε_i.
(ii) We consider $Y = \{N_{NRS_1}, ..., N_{NRS_n}\}$ and $X = \{TReqL_{CF_1}, ..., TReqL_{CF_n}\}$, X is obtained after observation of smart grid. The regression line equation is presented in Equation (11).

$$h(X_i) = \beta_0 + \beta_1 X_i + \varepsilon_i \tag{11}$$

where $h(X_i)$ refers to predicted number of integrated renewable sources for i^{th} observation. X_i is the predictor variable, it refers to total required load $TReqL_{CF_i}$ in a defined month i, β_0 and β_1 are regression coefficients, they represent y-intercept and slope of regression line respectively, ε_i is the independent normally irreducible error variable (Hong et al. 2010). The computed mean square error MSE between the function $h(X_i)$ and the computed numbers of new integrated renewable sources Y_i verifies the accuracy of results, i.e.,

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(h(X_i) - Y_i)^2 \tag{12}$$

(iii) We determine the new number of integrated sources for testing monthly values of $(N_{NRS_m}, TReqL_{CF_m})$, i.e.,

$$N_{NRS_m} = \beta_0 + \beta_1 * TReqL_{CF_m} + \varepsilon_i \tag{13}$$

(iv) Apply Simplex again to define the values of y_1, y_2 and y_3 for the new added renewable sources N_{NRS_m}.

3. Computing Threshold
This process allows to get the threshold number of renewable sources of SG, i.e.,

$$Thresh_{SG} = \frac{1}{n+m}\sum_{i=1}^{n+m} N_{NRS_i} \tag{14}$$

where N_{NRS_i} is the proposed new number of renewable sources to be added in i^{th} month and $n + m$ refers to the period of study in months. We choose the maximum value between N_{NRS_i} and $Thresh$ as a defined number of renewable sources to integrate.

III. CONTRIBUTION

In this section, we highlight the objective of this work that concentrates on enhancing recovery by improving best power availability and reducing the total waiting time of clients without electricity.

A. Smart Grid Reconfiguration Protocol

Every failure is categorized into dominant or equivalent in order to get a coverage source through a pertinent line or renewable source, or scheduled to be fixed by technical teams (Ziouche et al. 2019, Ben Meskina et al. 2018). We assume to propose a reconfiguration protocol that involves three phases as explained in Figure 1.

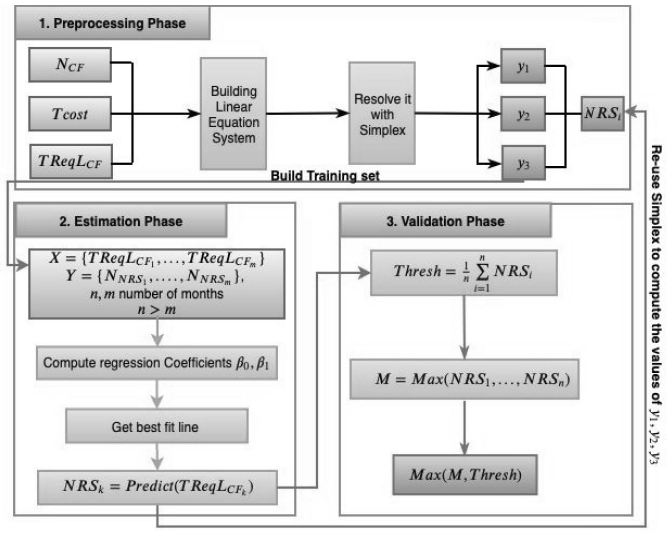


Figure 1: Smart Grid Reconfiguration.

1. Preprocessing Phase

A training dataset is built based on a system of equations using Linear Programming. Beside the given budget by power energy ministry $Tcost$, system receives as inputs critical failures number N_{CF} and total required load $TReqL_{CF}$. We apply Simplex method presented in section III.B.1 to compute the new added number of solar panels y_1, wind turbines y_2 and batteries y_3. These numbers are collected as $Y = \{N_{NRS_1}, .., N_{NRS_n}\}$.

2. Estimation Phase

In this phase, system is able to predict the accurate numbers of renewable sources to be integrated based on training datasets $X = \{TReqL_{CF_1}, .., TReqL_{CF_n}\}$ and $Y = \{N_{NRS_1}, .., N_{NRS_n}\}$. Thanks to the computed regression coefficients β_0 and β_1, we obtain the best fit line objective function that allows to predict the new number of added renewable sources of the tested total required load $TReqL_{CF_m}$ as explained in section III.B.2.

3. Validation Phase

System is able to compute the threshold value that corresponds to a given budget by choosing the maximum number of sources for the studied smart grid during a number of years, as reported in section III.B.3.

IV. EXPERIMENTATION & RESULTS

This section presents at first a study of the problematic, the proposed solution, performance evaluation metrics and the obtained results after applying this strategy on a real study.

A. Case Study

In order to explain the obtained gains and evaluation metrics, let us assume the power grid SG reported in

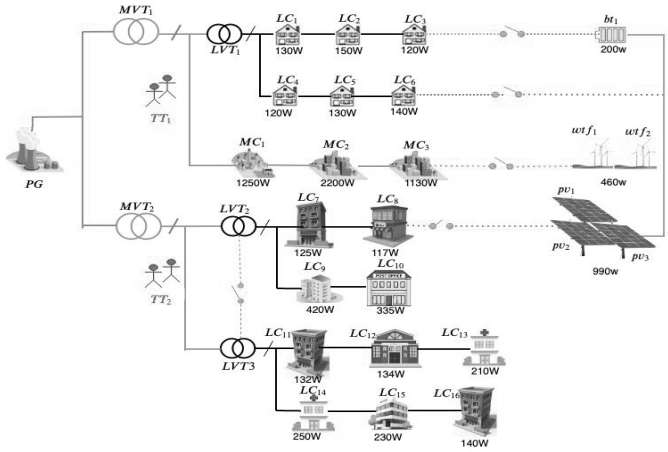


Figure 2: Small Smart Grid Running Example.

Figure 2. SG includes one generator PG, 5 transformers ($2MVT$, 3 LVT), 19 consumers (3 MC, 16 LC), 21 principal lines (5 MVL, 16 LVL) and 7 emergency lines (2 $eMVL$, 5 $eLVL$) with 6 renewable sources (1 bt, 2 wtf, 3 pv) and 2 technical teams. Let us consider

Table 1: The observed values of studied smart grid.

Month	Jan	Feb	Mar	Apr	May	Jun
Nf	10	9	11	8	12	13
N_{CF}	5	4	6	4	6	7
$TReqL_{CF}$	720	1600	1062	802	2062	2292
TWT_{CF}	$8h$	$10h$	$9h$	$7h$	$8h$	$12h$
TFT_{CF}	$7h$	$18h$	$14h$	$11h$	$12h$	$15h$
TRA_{RS}	100	250	0	17	115	230

that we supervise SG and define a set of global failures $SFL(SG)$ at every day, then, compute the mean number of failures per each month during six months, beside calculating the number of critical failures N_{CF}.

Problem: The insufficiency of energy available for recovery $TRA_{RS} < TReqL_{CF}$ and Highest waiting time of critical failures, beside least number of technical teams that cause late assistance as presented in Table 1.

The ministry proposes a budget $Tcost = 2950\$$ to buy a new number of renewable sources to integrate in SG. We apply the proposed reconfiguration to compute the added numbers of solar panels y_1, wind turbines y_2 and batteries y_3 for the smart grid SG, we suppose that: $N_1 = 3$, $N_2 = 2$, $N_3 = 1$, $P_{pv} = 710\$$, $P_{wtf} = 600\$$, $P_{bt} = 450\$$, $E_{pv} = 330W$, $E_{wtf} = 230W$, $E_{bt} = 200W$, $RT_{pv} = 3h$, $RT_{wtf} = 2h$ and $RT_{bt} = 3h$.

Phase 1: we implement the first iteration table to compute the values of y_1, y_2 and y_3 as explained in Table 2. In every iteration, we choose the pivot line and column, we keep computing till getting the values of y_1, y_2 and y_3. We apply the same algorithm at the first three months to generate a training set and get of the first

three months as presented in Table 3, beside the total satisfied load TSL_{CF} and total remaining load TRL_{CF}.

Table 2: Fist Iteration Table of the first month

Base Vars	y_1	y_2	y_3	e_1	e_2	e_3	b_i
e_1	710	600	450	1	0	0	2950
e_2	3	2	3	0	1	0	15
e_3	330	230	200	0	0	1	620
N_{NRS}	1	1	1	0	0	0	0

Phase 2: we suppose that $X = [720, 1600, 1062]$ and $Y = [3, 4, 4]$, we compute the values of β_0 and β_1. We find that $\beta_0 = 2.50000677502$ and $\beta_1 = 0.0010348845875$. Hence, the best fit line function is $h(X_i) = 2.5 + 0.001 * X_i + 0.0001$. We predict the values of new added renewable sources for the last three months, and we re-use Simplex to define the values of y_1, y_2, y_3, TRL_{CF} and TSL_{CF} as presented in Table 3.

Table 3: Obtained results due to Simplex Method.

Month	y_1	y_2	y_3	N_{NRS}	TRL_{CF}	TSL_{CF}
Jan	0	1	2	3	$0W$	$720W$
Feb	2	0	2	4	$290W$	$1310W$
Mar	2	0	2	4	$2W$	$1060W$
Apr	2	0	2	4	$0W$	$802W$
Mai	2	1	2	5	$917W$	$1060W$
Jun	2	1	2	5	$1030W$	$1060W$

Phase 3: system permits to get the threshold value, where: $Thresh = 5$ and $y1 = 2, y_2 = 1, y_3 = 2$ and we determine the required load satisfied by the number of sources integrated in SG: $TSL_{CF} = 1320W$ and $TRTA_{RS} = 14h$ with $N_{RF} = 8$.

B. Performance Evaluation

This work addresses a strategy that permits to predict the exact number of new renewable sources to be added into a defined smart grid. So, we define two metrics that are power availability and total waiting time rates.

1. Power Availability Rate

-Denoted by σ_{SG}-, it represents the power availability guaranteed due to applied approach to enhance recovery. Total availability presents the rate of new number of recovered failures compared with total number of failures, i.e.,

$$\sigma_{SG} = \frac{N_{RF} + N_{RF_{init}}}{Nf} \tag{15}$$

where N_{RF} is the new number of recovered failures and $N_{RF_{init}}$ is the initial number of recovered failures.

2. Total Waiting Time Rate

- Denoted by ρ_{SG}-, it permits to determine the gain in term of time before and after applying the proposed forecasting approach, i.e.,

$$\rho_{SG} = \frac{TWT_{end}(SG)}{TWT_{init}(SG)} \tag{16}$$

where $TWT_{end}(SG)$ and $TWT_{init}(SG)$ refers to the total waiting time after and before inserting the new recovery sources.

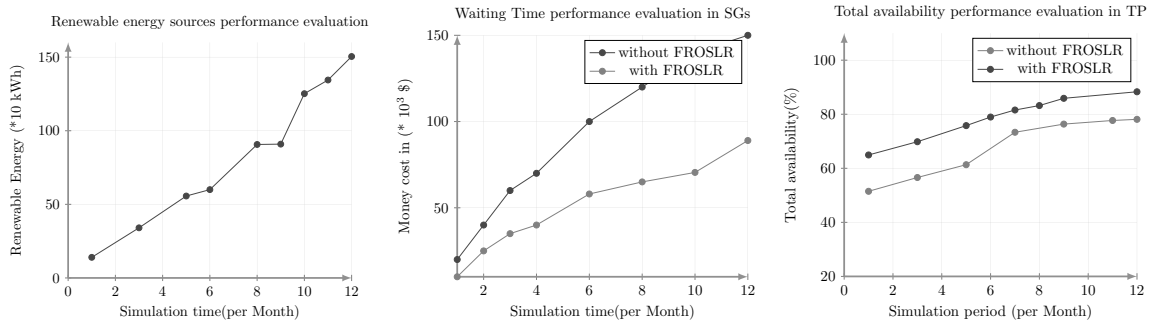
Table 4: Investigated Smart Grids.

SG	N_{NRS}	N_{RF}	TWT_{CF}	TSL_{RF}	$\sigma(SG)$
SG_1	3	5	$6h$	860W	85.29%
SG_2	4	7	$12h$	1190W	79.51%
SG_3	7	10	$21h$	2050W	78%
SG_4	8	15	$24h$	2510W	81.11%

In order to improve the performance of the proposed contribution, we have applied it on four real meshed power networks of the Tunisian electric grids. They are with variant sizes where SG_1, SG_2 are two small grids, SG_3 is medium and SG_4 is large, they include different numbers of components and lines. They are simulated using -FROSLR- simulator for two years. Their structures are shown in Table 5, where we have the number of consumers $|CONS|$, number of pertinent emergency lines $|EL|$, number of renewable sources $|RS|$, number of recovered failures N_{RF} and non recovered failures N_{NRF}, beside the initial total power availability $\sigma_{init}(SG)$. To test and validate the implemented tool, we have injected the strategy on each studied electric grid, we have computed for every smart grid the total required load to be satisfied by new added renewable sources in order to enhance the power availability during recovery. Table 4 shows the obtained gains after adding a number of renewable sources, we note that there exists a remarked increasing in term of total power availability rate $\sigma(SG)$. Furthermore, this strategy permits to reduce the total waiting time for the assistance of technical teams to repair the outages and interruptions. Figure 3a shows the enhancement of use of renewable energy in the Tunisian power electric grids per one year into $1500kWh$, where it represents the exploiting of the renewable energy sources increasing that permits to recover large number of failures. Figure 3b shows the obtained gains in term of financial cost, where the figure shows a remarkable gain in term of money. It permits to provide more than 140.000 $ per year due to proposed -FROSLR- tool. Figure 3c presents the total power availability enhancement in smart grid SG_4, the supervision shows that there exists an enhancement from 10% to 20%. The measures of performance rates of SG_1 are significantly rises from 70.58% to 85.29%.

Table 5: Investigated Smart Grids in Experimental Study.

SG	$\|CONS\|$	$\|EL\|$	$\|RS\|$	N_f	$N_{RF_{init}}$	$TReqL_{CF}$	$Tcost$	TWT_{CF}	$\sigma_{init}(SG)$
SG_1	31 LC & $3MC$	12	3	34	24	1400	1950$	10h	70.58%
SG_2	75 LC & 8 MC	25	5	83	59	1640	2780$	27h	71.08%
SG_3	122 LC & 28 MC	54	7	150	107	3230	4750$	1day2h	71.33%
SG_4	142 LC & 38 MC	65	10	180	131	4067	5470$	2days	72.78%



(a) Renewable energy increasing. (b) Gain in term of money cost. (c) Total availability enhancement.

Figure 3: Performance Evaluation Enhancement based on FROSLR Simulator.

V. CONCLUSION

The problem of the least number of technical teams that intervene in failures repairing, and the increasing of economic losses per year, propose new challenges that search for a reconfiguration that permits to guarantee a best power availability, reduce total waiting time and recorded losses based on predicting the new number of renewable sources to be integrated into a defined grid. In this paper, we have developed a tool called -FROSLR- for smart grid failures number optimization and maximum load required for recovery using Simplex and Linear Regression. We briefly present: (i) smart grid failures management, (ii) failures classification according to recovery sources of critical and non-critical, (iii) determine the number of renewable sources to integrate due to a given budget by ministry, and (iv) re-compute the rates of total waiting time and total power availability. In the future works, we will discuss the heuristic of optimizing the number of technical teams that assist in maintenance and estimation of scheduling tables times.

REFERENCES

Abidi,M. G.,Ben Smida,M., Khalgui,M., Li, Z., and Wu,N.(2017). "Multi- agent oriented solution for forecasting-based control strategy with load priority of microgrids." *Electric Power SR.* 152, 411-423.

Ben Meskina, S., Doggaz,N., Khalgui, M. and Li, Z. (2017). "Multi-Agent Framework for Smart Grid Recovery". *IEEE SMC.* 47(7) 1284-1300.

Ben Meskina, S., Doggaz,N., Khalgui, M. and Li, Z. (2018). "Reconfiguration-based Methodology for Improving Recovery Performance of Faults in Smart Gids". *Elsevier.* 454, 73-95.

Chowdhury, Md. M., Dass,P., Nygard, K. E., and Khiabani, V. (2016). "Optimization Modeling in a Smart Grid." *Research Gate.* ICWN'15, 85-90.

Burkard, Rainer. E., Dell'Amico, M., and Martello, S. (2009). "A genuinely polynomial simplex algorithm for the assignment problem." *Elsevier.* 45(2), 93-115.

Hong, T., Gui,M., Baran, M. E. and Lee Willis, H. (2010). "Modelling and forecasting hourly electric load by multiple linear regression with interactions". *IEEE PES General Meeting.* 1-18.

Nefedov,E., Sierla, S., and Vyatkin,V. (2018). "Internet of Energy Approach for Sustainable Use of Electric Vehicles as Energy Storage of Prosumer Buildings." *Energies.* 11, 2165

Ogryczak, W. (1988). "The simplex method is not always well behaved. Linear Algebra and Its Applications." *Springer.*, New York, NY 10010.

Yuce, B., Mourshed, M., and Rezgui, Y. (2017). "A Smart Forecasting Approach to District Energy Management." *Energies.* 10, 1073.

ZAFAR, S., NAWAZ, K. , NAQVI, S, and MALIK, T.N. (2013). "Integration of Renewable Energy Sources In Smart Grid: A Review." *Nucleus* 50(4), 311-327

Ziouche, L., Ben Meskina, S., Khalgui, M., Kahloul, L. and Li, Z. (2019). "Smart Grid Re-building based on Cloud Computing Architecture. " *IEEE SMC*, 1-8.

MULTI AGENT SYSTEM-BASED APPROACH FOR ENHANCING CYBER-PHYSICAL SECURITY IN SMART GRIDS

Ilyes Naidji
Faculty of Mathematical, Physical and Natural Sciences,
University of Tunis-El Manar, Tunisia.
Email: ilyes.naidji@fst.utm.tn

Moncef Ben Smida
LSA Lab, Tunisia Polytechnic School,
University of Carthage, Tunisia.
Email: moncef.bensmida@gmail.com

Mohamed Khalgui
National Institute of Applied Sciences and Technology,
University of Carthage, Tunisia.
Email: khalgui.mohamed@gmail.com

Abdelmalik Bachir
LESIA Lab, Computer Science Department,
Biskra University, Algeria.
Email: a.bachir@univ-biskra.dz

KEYWORDS
Smart grid, cyber physical attacks, agents, security, physical faults, protection

ABSTRACT

Smart grids are critical cyber-physical power systems that are characterized by reliable and safe operation against physical faults and cyber attacks. Smart grid systems have to be developed for the protection from cyber-physical threats by taking the necessary countermeasures. To address this problem, this paper proposes a unified framework that can detect, identify and reconfigure the smart grid from cyber attacks as well as physical faults to recover the system operation. The proposed framework is based on sensitivity analysis which can estimate the system state in efficient way. The framework has a multi-agent-based architecture that consists of several types of agents with well-defined behaviors which cooperate to fulfill the goal of cyber-physical system (CPS) protection against cyber-physical attacks. We evaluate the performance of the proposed approach using reliability and economy metrics for a cyber-physical system aimed at monitoring a smart grid.

INTRODUCTION

A smart grid is a complex network which comes with many security issues and threats. Real world incidents as cascading failures or malicious actions causing several blackouts have demonstrated the incapacity of power grids to maintain a reliable service in such cases. Thus, the need for a reliable and self-healing network which ensures the security against physical faults and cyber attacks. Existing solutions for smart grid security and stability are mainly focused on state estimation techniques for the set of network buses (Mosbah and El-Hawary 2018). These solutions are static and suffer from several limitations. The majority of the existing control and monitoring tools in energy management systems are based on steady-state of the power system, which cannot take into account the system dynamics (Kamwa et al. 2016). This limitation is due to the energy management system functions that rely on the Supervisory Control and Data Acquisition (SCADA) systems, which have slow scan rates without time stamps. Thus, state estimation is updated every minutes. Furthermore, the control schemes associated with generators and Flexible AC Transmission Systems (FACTS) devices are based on local measurements without considering the measurements of the other existing generators in the network.
For this reason, the development of security solutions by exploiting the dynamics of the power system should be addressed. From static to dynamic state estimation (Tebianian and Jeyasurya 2013), several studies have been carried out to address the security and stability of the smart grid by taking into account the dynamics of the power system (Bartolomey and Semenenko 2017). An attack detection algorithm is proposed in (Zhou et al. 2018) which can effectively identify common types and numbers of network attacks. However, this study is much quantitative than qualitative, i.e., the algorithm detects the number of attacks and just common type of attacks without specifying real solutions.
Other studies consider game theoretic approaches for cyber security problems in cyber-physical systems by modeling the problem as a game between attacker and defender. The work in (Guo et al. 2018) proposes a game-theoretic approach for cyber-physical systems to analyze the interactions between an attacker and a defender and develops effective defense strategies. A behavioral framework of cognitive hierarchy theory is proposed where cyber physical security of smart grid is addressed in (Sanjab and Saad 2016). The framework analyzes the interactions between one defender and one attacker that can have different computational capabilities and different levels of knowledge of the system.
Thus, previous work has focused on strategic behavior of players who have full knowledge of the payoff of the other players and full knowledge of system state. Furthermore, most of these studies have modeled the player interactions at the decision-making level rather than for-

mulating a cyber-physical system model. Thus, there is a need to study concrete models of incorporating security in smart grid controllers.

In this paper, the protection against cyber-physical attacks is addressed where cyber attacks are in the form of integrity attacks where it is assumed that Phasor Mesurement Units (PMUs) are attacked . Consequently, the measurements of PMU are altered resulting in unintentional control commands that lead to a significant disturbances in the system. This may conduct an operational line or generator to be out-of-service, even if there is no physical fault. For this reason, efficient algorithms are proposed to deal with cyber-physical attacks. These algorithms use the results of a sensitivity analysis which is based on four indices that are voltage magnitude measurement, line current measurement, power generation measurement, and circuit breaker status. The proposed algorithms are executed by an intelligent multi agent system (MAS) based relay protection. The MAS performs sensitivity analysis in order to detect an attack and identify its nature, i.e., distinguish if it is physical or cyber to take the necessary countermeasures and restore the system in its normal state. The contributions of this paper are:

- The proposal of a dynamic protection scheme to ensure both cyber and physical security of the smart grid by updating the system state in real time.
- The estimation of the system state in an efficient way using a smart sensitivity analysis.
- The preservation of the system state security by keeping the network constraints to an acceptable level.

An illustrative case study based on the IEEE 9 bus system is provided to validate the effectiveness of the proposed framework.

This paper is organized as follows: Section 2 presents the system model. Section 3 formulates the problem of cyber physical security in smart grids. Section 4 gives the proposed methodology for solving the cyber-physical security problem. Section 5 shows the simulation results and finally Section 6 concludes this paper.

SYSTEM MODEL

The smart grid system is divided into multiple subsystems. The physical connection is provided by transmission power lines while communication is provided by agents. A group of agents is developed for an individual subsystem. Each two subsystems are connected physically as well as through communication provided by agents. In this paper, it is assumed that the smart grid already has a communication network.

PMU-based Relay Infrastructure

Phasor Measurement Units (PMUs) are currently used in smart grids to exchange information in real time about voltage, current and status of circuit breakers (CBs). With the apparition of digital relays that are PMU-based, the agent-based protection relay adds a new feature to the smart grid security. Fig. 1 illustrates a typical node in the smart grid network where each bus is equipped with a protective relay. The smart grid

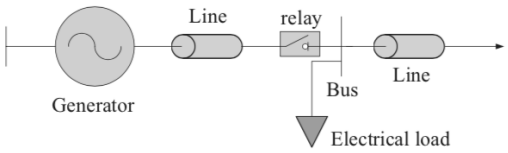


Figure 1: Typical node in smart grid

has an agent-based protection relay architecture where agents are integrated with protection relays. The agents can play the role of the bus controller where they can manage the generation, load and the protective relays to have a reliable energy management.

Energy Management System

The energy management system (EMS) consists of different applications that control and monitor the state of the smart grid. The EMS takes decisions based on algorithms that visualize and process the measurements collected by different PMUs. Our EMS consists of a software based on Matpower package. The proposed EMS is an optimization of the EMS developped in (Naidji et al. 2018) and (Naidji. et al. 2019). The proposed EMS calculates the optimal power flow to have the minimum generation cost while supplying all the loads in the network.

PROBLEM FORMULATION

Cyber physical attacks are in the form of additive sensor values causing the loss of synchronization of generators and leading to an imbalance between demand and supply. Cyber Security requires the control of integrity of data and knowledge of digital control systems. Knowing the protocol of digital control systems, an attacker can manipulate the critical data to inject a malicious information. Since binary status signals are used for the control of circuit breaker (CB) operation, therefore, a successful cyber attack may results in unscheduled opening of circuit breakers (CBs). A malicious change of the control signals can alter the behaviors of the system causing ultimate losses.

Cyber Attack Model

The information obtained by the PMU agent may be sent via unsecured links. When the information experiences a cyber attack, it could be modified by a fake value. This results in wrong or inaccurate meter data (e.x., voltage or line current...), such that, the input entering the direct control agent is different from the real values which can alter the control commands. This change in meter data values can result in loss of synchronization between the generators leading to an imbalance between demand and supply. Furthermore, during a cyber attack, an incorrect countermeasures can destabilize the system causing significant damages to the system operator.

False Data Injection Attack
Let $\mathbf{z_i}$ denotes the vector of measurements of given by the PMU. Let us assume that each state of subsystem can be affected by a cyber attack. Let $\mathbf{a_k}$ denotes the vector of attack which is generated in the form of an additive information. The vector of attack is added to the original measurements of the PMU as follows:

$$\mathbf{z_i} = \mathbf{z_i} + \mathbf{a_k} \tag{1}$$

Alert Manipulation Attack
Furthermore, a cyber attack can be in the form of alert manipulation attack. In this scenario, the attacker aims to deceive the energy management system through launching false alerts like generator tripping or transmission line tripping. As a result of this alert, the EMS will make a decision based on the manipulated alert. To demonstrate this scenario, we introduce several false alerts to the energy management system with and without using the proposed protection scheme to detect, identify and reconfigure the network.

DISTRIBUTED MAS ARCHITECTURE

Modern smart grids combine physical with cyber system infrastructure where information and communication technologies (ICT) are embedded in power generation, transmission and distribution to consumer appliances. We present a distributed multi agent based architecture for Smart Grid that provides a feasible solution for identifying cyber attacks and physical faults. The proposed framework distinguish between physical faults and cyber attacks and considers physical and cyber activities in the smart grid. The proposed multi-agent model is a collection of agents that cooperate by exchanging information to control, monitor and protect the smart grid system against cyber attacks.

Fig. 2 shows the considered network which is the IEEE 9 bus system with an agent-based protection relay architecture where agents are integrated with protection relays.

The network is divided in multiple sub-systems where each subsystem is monitored by three type of agents that are PMU agent, Detection and Identification agent, and Dynamic Control agent. The PMUs continuously provide the measurement of the current and voltage. Furthermore, PMUs provide a set of circuit breaker (CB) signal status to support the decision making of relays. These protection relay agents are located in different nodes of the network. The role of these agents is determined by software. The agents use CB status signals to execute their control actions during disturbances.

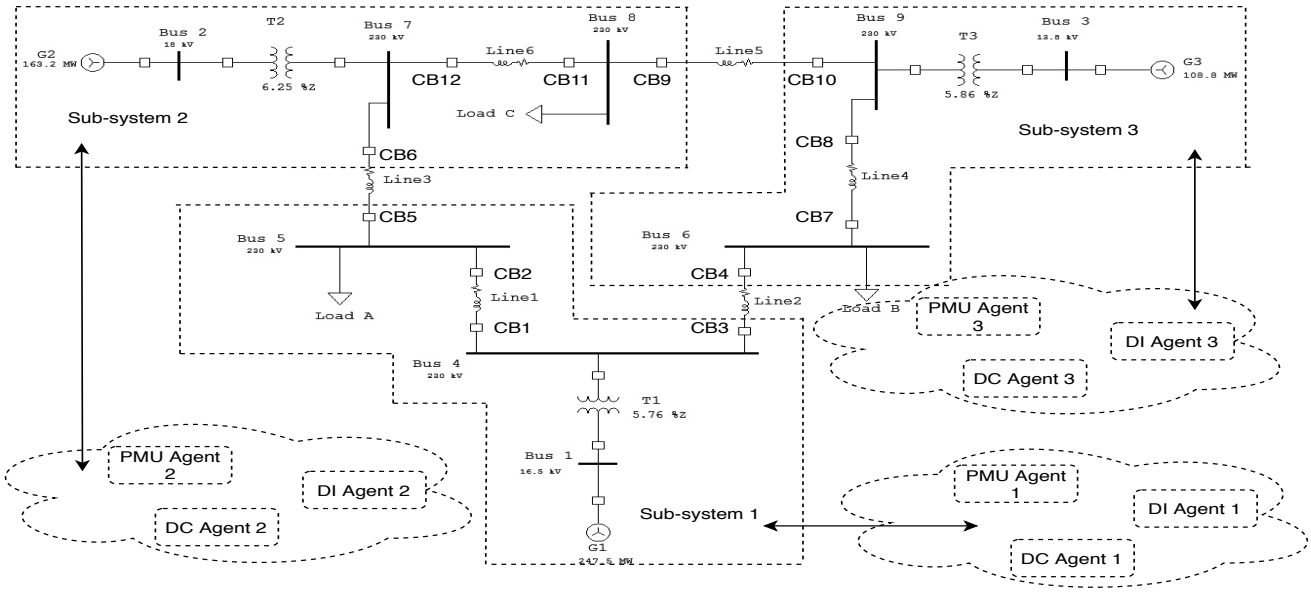


Figure 2: IEEE 9 bus system

Role of Agents

PMU Agent
The PMU agent obtains the information from phasor measurement units (PMUs) of the measured voltage $\mathbf{z_V}$, line current $\mathbf{z_I}$ and circuit breaker status $\mathbf{z_{CB}}$ to monitor the network topology. Furthermore, the PMU agent sends information about measurement taken from PMU to the detection and identification agents.

$$\mathbf{z_I} = \left[I_1, ..., I_{|\mathbf{z_I}|}\right] \tag{2}$$

$$\mathbf{z_V} = \left[V_1, ..., V_{|\mathbf{z_V}|}\right] \tag{3}$$

$$\mathbf{z_{CB}} = \left[CB_1, ..., CB_{|\mathbf{z_{CB}}|}\right] \tag{4}$$

where I and V are the measured phase currents and voltages. The PMU agent records the status of the lines in a status matrix. The rows and columns of the matrix represents the set of network buses.
This matrix are filled with binary variables z_s where the status '1" indicates the presence of flow between buses, i.e., the line is connected, and '0" indicates the absence of flow between buses, i.e., the line is disconnected. The status matrix is given by

$$\mathbf{Z_S} = \begin{bmatrix} 0 & 1 & 1 & ... & 1 \\ 1 & 0 & 1 & ... & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & ... & 0 \end{bmatrix} \tag{5}$$

Detection and Identification Agent
The DI agent detects if there is a deviation in line current or voltage magnitudes. In addition, it checks the status of circuit breakers and compare them with the previous states of the system. If there is a change in status, then the agent performs some tests to detects if this change is normal or anomalous. After that, if the result of tests shows that there is an abnormal change, the agent identifies if it is a physical fault or a cyber attack. To accomplish this task, the DI agent performs sensitivity analysis to calculate a deviation vectors for: line current denoted by $\mathbf{I_D}$, voltage denoted by $\mathbf{V_D}$ circuit breaker denoted by $\mathbf{CB_D}$, and power generation denoted by $\mathbf{PG_D}$, i.e.,

$$I_D = \begin{cases} 1, & \text{if } I - I^{prev} > t_l, \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

$$V_D = \begin{cases} 1, & \text{if } V - V^{prev} > t_V, \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

$$CB_D = \begin{cases} 0, & \text{if } CB - CB^{prev} = 0, \\ 1, & \text{otherwise} \end{cases} \tag{8}$$

$$PG_D = \begin{cases} 0, & \text{if } PG - PG^{prev} > t_P, \\ 1, & \text{otherwise} \end{cases} \tag{9}$$

If the DI agent receives a notification of a physical fault, then DI agent confirms this fault by checking the measurements of PMU, i.e., line current, voltage and CB status. A threshold value for current, voltage and power generation that are t_I, t_V and t_P respectively are set for each DI agent. Thus, during a physical fault in a particular line or bus, DI agent can detect it by measuring the fault current and voltage. That means when current flows above the predefined threshold or voltage exceeds its allowable interval, i.e., the DI agent realizes a physical fault in the considered bus in the system. Alg. 1 illustrates the role of the detection and identification agent in case of false data injection attacks while Alg. 2 illustrates the role of the detection and identification agent in case of alert manipulation attacks.

Algorithm 1: Detection and Identification for False Data Injection Attack

Input: PMU measurements
Output: system state (ss)
if $I_D = 1$ *and* $z_s = 0$ **then**
  $ss = line\ fault$
else
  $ss = cyber\ attack$
end
if $CB_D = 1$ *and* $PG = 0$ **then**
  $ss = generator\ fault$
else
  $ss = cyber\ attack$
end
if $V_D = 1$ *and* $PG_D = 1$ **then**
  $ss = voltage\ fault$
else
  $ss = cyber\ attack$
end
return ss;

Once the physical fault is identified, the DI agent sends information to the DC agent to coordinate the agent-based protection relays using the CB status signals. However, when the measurements of PMU disagree that there is a physical fault, a cyber attack is confirmed while the fault is denied. The DI agent sends information about system state (SS) to the DC agent to confirm that there is a cyber attack or physical fault.

Dynamic Control Agent
The dynamic control (DC) agent receives information, i.e., system state (SS) from DI agent about the encountered problem, i.e., the considered physical fault or cyber attack reconfigure the system by taking the necessary countermeasures. Alg. 3 illustrates the role of the dynamic control agent.
The algorithm takes the system state as an input to take the necessary countermeasures. If the system is undergoing a cyber attack, the altered values are replaced by

Algorithm 2: Detection and Identification for Alert Manipulation Attack

Input: PMU measurements
Output: system state (ss)
if *transmission line tripping alert* **then**
  if $I_D = 1$ *and* $PG_D = 1$ **then**
    $ss = linefault$
  else
    $ss = cyberattack$
  end
else if *generator tripping alert* **then**
  if $CB_D = 1$ *and* $PG = 0$ **then**
    $ss = generatorfault$
  else
    $ss = cyberattack$
  end
else if *voltage alert* **then**
  if $V_D = 1$ *and* $PG_D = 1$ **then**
    $ss = voltagefault$
  else
    $ss = cyberattack$
  end
return ss;

Algorithm 3: Reconfiguration

Input: system state (ss)
Output: necessary countermeasures
if *ss=cyber attack* **then**
  if $\mathbf{V_D} = 1$ **then**
    $V = V^{prev}$
  else if $I_D = 1$ **then**
    $I = I^{prev}$
  else if $CB_D = 1$ **then**
    $CB = CB^{prev}$
else if *ss=physical fault* **then**
  Isolate fault;
  Repair system;
  Reconfigure system operation;

the precedent values and no countermeasures are taken, else if the system is undergoing a physical fault, the considered fault is isolated, then the system is repaired and finally the system operation is reconfigured.

EXPERIMENTATION

In this section, we report on the performed attack scenarios and collected results. We perform some simulations to show the impact of cyber-physical attacks on the IEEE 9 bus system.

The attacker introduces a false alert manipulation attack which is transmission line tripping (line between bus 1 and 4) to deceive the energy management system. The attack impact on power generation is illustrated in Fig. 3. This manipulation raised a false alarm at the energy management system which leads to an undesirable countermeasures. As Fig. 3 shows, the power generation of generator 2 and 3 is increased to compensate the presumed power shortage caused by the line tripping. However, in fact generator 1 is operational leading to an extra power surplus which increases power losses and generation cost. After, one hour on extra generation, the EMS realizes that there is no generation tripping, so generator 2 and 3 stop the extra generation and return to their typical generation level. However, this extra generation costs a lot to the smart grid operator.

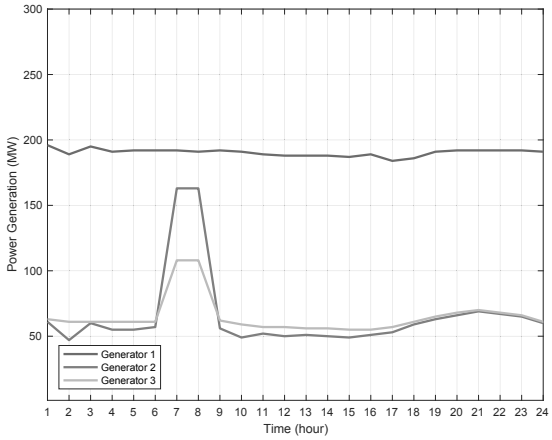


Figure 3: Impact of alert manipulation attack on power generation

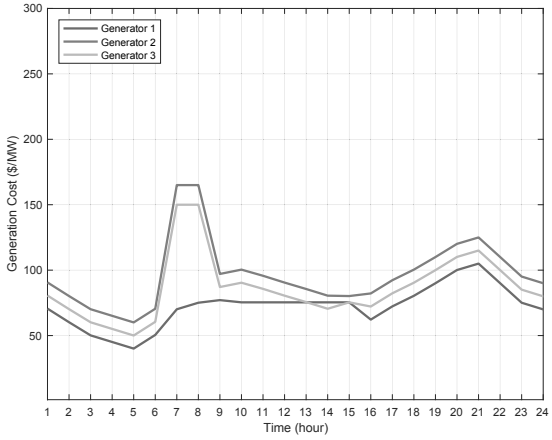


Figure 4: Impact of alert manipulation attack on generation cost

As shown in Fig. 4, the generation cost is increased during peak hours, however, due to nature of cost function which is a quadratic function, the cost is increased during cyber attacks with 51% in generator 2 and with 43%

in generator 3. Furthermore, this increase is caused by the use of generation reserves which have a higher cost. For this reason, the need for using an efficient protection scheme for the security of the smart become apparent. The proposed protection scheme is based on a distributed Multi Agent System (MAS) architecture that cooperate to fulfill the goal of smart grid security. After the introduction of an alert, the agents perform three steps that are attack detection, identification and reconfiguration.

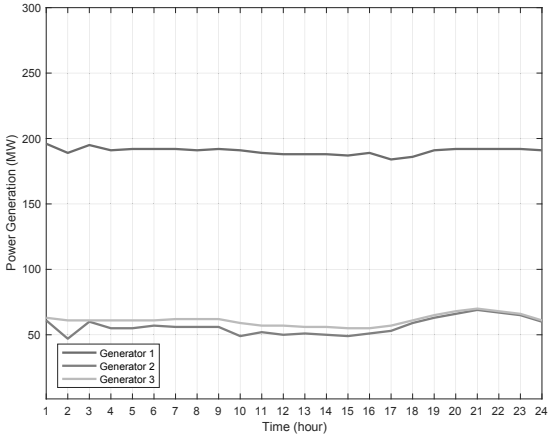


Figure 5: Power generation with the proposed protection scheme

Fig. 5 shows the non impact of cyber attack on power generation, thanks to the fast detection, identification and reconfiguration algorithms that ensure the security of the smart grid. After the introduction of a cyber attack, the proposed MAS detects and identifies the alert, after the cyber attack identification, no countermeasures will be taken, such that, no extra generation will be executed, thus no extra generation cost will be added.

CONCLUSION

A dynamic protection scheme for smart grids is proposed. The proposed protection scheme secures the network from cyber attacks as well as physical faults by the coordination of agents to detect and identify attacks and reconfigure the system operation. The detection and identification is performed by sensitivity analysis which estimates the system state in an efficient way. To ensure the smart grid protection against cyber-physical attacks, a reconfiguration algorithm is executed to restore the system operation in a secure way. As a perspective, other identification techniques as well as new cyber-physical attacks in smart grids will be addressed in a future work.

REFERENCES

Bartolomey P. and Semenenko S., 2017. *Super-accelerated power systems power flow and state estimation calculations within the WAMS environment.* In *2017 14th International Conference on Engineering of Modern Electric Systems (EMES)*. 55–58. doi: 10.1109/EMES.2017.7980380.

Guo Y.; Gong Y.; Njilla L.L.; and Kamhoua C.A., 2018. *A stochastic game approach to cyber-physical security with applications to smart grid.* In *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. 33–38. doi: 10.1109/INFCOMW.2018.8406833.

Kamwa I.; Geoffroy L.; Samantaray S.R.; and Jain A., 2016. *Synchrophasors data analytics framework for power grid control and dynamic stability monitoring. IET Eng Technol Ref*, 1–22.

Mosbah H. and El-Hawary M., 2018. *Power System Static State Estimation Using Modified Stochastic Fractal Search Technique.* In *2018 IEEE Canadian Conference on Electrical Computer Engineering (CCECE)*. ISSN 2576-7046, 1–4. doi:10.1109/CCECE.2018.8447826.

Naidji. I.; Mosbahi. O.; Khalgui. M.; and Bachir. A., 2019. *Cooperative Energy Management Software for Networked Microgrids.* In *Proceedings of the 14th International Conference on Software Technologies - Volume 1: ICSOFT,*. INSTICC, SciTePress. ISBN 978-989-758-379-7, 428–438. doi:10.5220/0007965604280438.

Naidji I.; Smida M.B.; Khalgui M.; and Bachir. A., 2018. *Non cooperative game theoretic approach for residential energy management in smart grid.* In *Proceedings of the the 32nd Annual European Simulation and Modelling Conference*. ETI, EUROSIS, 164–170.

Sanjab A. and Saad W., 2016. *On bounded rationality in cyber-physical systems security: Game-theoretic analysis with application to smart grid protection.* In *2016 Joint Workshop on Cyber-Physical Security and Resilience in Smart Grids (CPSR-SG)*. IEEE, 1–6.

Tebianian H. and Jeyasurya B., 2013. *Dynamic state estimation in power systems using Kalman filters.* In *2013 IEEE Electrical Power Energy Conference*. 1–5. doi:10.1109/EPEC.2013.6802979.

Zhou C.; Wang Z.; Huang W.; and Guo Y., 2018. *Research on Network Security Attack Detection Algorithm in Smart Grid System.* In *2018 IEEE International Conference of Safety Produce Informatization (IICSPI)*. 261–264. doi:10.1109/IICSPI.2018.8690458.

ENVIRONMENTAL AND POPULATION DYNAMICS SIMULATION

STOCHASTIC MODEL OF THE TIME SERIES OF THE AVERAGE DAILY BIOCLIMATIC INDEX OF SEVERITY OF CLIMATIC REGIME

Nina A. Kargapolova
Institute of Computational Mathematics and Mathematical Geophysics SB RAS
Pr. Lavrent'eva 6
630090 Novosibirsk
Russia

Novosibirsk State University
Pirogov St. 1
630090 Novosibirsk
Russia

E-mail: nkargapolova@gmail.com

KEYWORDS
Stochastic Simulation, Time-series Analysis, Non-stationary Random Process, Bioclimatic Index of Severity of Climatic Regime, Climate Change.

ABSTRACT

In this paper, a numerical stochastic model of the time series of the average daily bioclimatic index of severity of climatic regime is proposed and validated. This model is based on an assumption that real weather processes are non-stationary random processes on a year-long interval. In this assumption, the model takes into account the seasonal variation of the real meteorological processes. The input parameters of the model are determined from the data of long-term real observations at weather stations. It is shown that the trajectories of the model proposed are close in their statistical properties to the real time series of the bioclimatic index under consideration. The results related to studying the influence of a climate change on the time series of the average daily bioclimatic index of severity of climatic regime are given.

INTRODUCTION

The weather conditions, including air temperature and relative humidity, wind speed and atmospheric pressure, have a significant impact on the state of health of a human being and his / her ability to work. To describe the combined effects of various meteorological parameters, different climatic indicators and weather indices are used. Several of them (for example, the heat index and the apparent temperature) are used to describe the effects of high temperatures and high relative humidity in the summer season (Anderson et al. 2013, Steadman 1994). Other indices, such as the Siple index and the Hill wind chill index, are used to characterize the effects of low temperatures and high winds during the cold period of a year (Hill et al. 1958, Osczevski and Bluestein 2005). One of the most universal bioclimatic indicators is the bioclimatic index of severity of climatic regime (Belkin and Poltorak 1983, Belkin et al. 1989).

At present, to study the properties of the time series of bioclimatic indices, the two approaches are mainly used. The first approach is the statistical one. In the framework of this approach, data from real observations are analyzed, see, for example, (Kershaw and Millward 2012, Revich and Shaposhnikov 2016, Shartova et al. 2018). The second approach to studying the properties of the time series of bioclimatic indices is based on the use of hydrodynamic models of atmospheric processes (Gosling et al. 2009, Ohashi et al. 2014).

In 2018, the author of the paper together with colleagues at the Institute of Computational Mathematics and Mathematical Geophysics (Novosibirsk, Russia) and at the Voeikov Main Geophysical Observatory (St. Petersburg, Russia) began the development of a stochastic approach to studying and simulation of the time series of bioclimatic indices. For short time intervals (about 10–12 days), models of the high-resolution time series of the heat index, the enthalpy of the humid air and the bioclimatic index of severity of climatic regime (BISCR) were constructed (Akenteva et al. 2019, Kargapolova, 2019). These models take into account the daily variations of the real meteorological time series.

In this paper, a stochastic model of the time series of the average daily bioclimatic index of severity of climatic regime is proposed and validated. It is shown that the trajectories of the model proposed are close in their statistical properties to the real time series of the BISCR. An example of an application of the model to studying the influence of a climate change on the BISCR is given.

BIOCLIMATIC INDEX OF SEVERITY OF CLIMATIC REGIME

The bioclimatic index of severity of climatic regime was proposed by V.S. Belkin, a Soviet researcher (Belkin and Poltorak 1983, Belkin et al. 1989). According to (Finaev 2004), the BISCR is an integral indicator of the degree to bioclimatic discomfort in various types of the vital activity. The BISCR depends on a numver of meteorological and physiographic parameters: air temperature (measured in degrees Celsius), atmospheric pressure (in hectopascals), wind speed modulus (in meters per second), relative humidity of the air (measured in percentage), and elevation of the terrain under consideration above the sea level (in meters). The BISCR is defined as

$$B = \frac{\tilde{T}\left(P - 266\right)\left(1 - 0.02V\right)}{75\tilde{R}\tilde{S}},$$

where $\tilde{T}, \tilde{R}, \tilde{S}$ are the temperature, humidity and radiation coefficients, P, V are the atmospheric pressure and wind

speed modulus, respectively, (Belkin et al. 2016). The coefficients are determined by the following formulas:

$$\tilde{T} = \begin{cases} 1 - 0.0089(22 - T), & T < 22, \\ 1 - 0.0263(T - 22), & T \geq 22, \end{cases}$$

$$\tilde{R} = \begin{cases} 1 + 0.6\dfrac{50 - R}{100}, & R < 50, \\ 1 + 0.6\dfrac{R - 50}{100}, & R \geq 50, \end{cases}$$

$$\tilde{S} = \begin{cases} 1, & H < 2000, \\ 1 + 0.45\dfrac{H - 2000}{1000}, & H \geq 2000. \end{cases}$$

Here T, R are air temperature and relative humidity and H is the elevation of the terrain above the sea level.

The values ot the BISCR belongs to the interval $[0;10]$. If $B \in [8;10]$, then the weather conditions are comfortable. The values of the BISCR belonging to the intervals $[7;8)$ and $[6;7)$ denot the two levels of relative comfort. The weather conditions, where $B \in [5;6)$ or $B \in [4;5)$, cause varying degrees of compensable discomfort to a human being, and $B \in [0;4]$ designates non-compensable discomfort.

STOCHASTIC MODEL

Let us describe a stochastic model of the time series $\vec{B} = (B_1, B_2, \ldots, B_8, B_9, \ldots, B_{365})$ of the average daily BISCR on a yearlong interval.

In (Akenteva et al. 2019) an approach to simulating the high-resolution time series of the BISCR as functions of the joint mereorological time series was proposed. This approach is based on the assumption that the meteorological time series are periodically correlated random processes at short time intervals. At a year-long time interval, the time series are non-stationary, thus resulting in complication of the model and in a heavy increase in its computational costs.

In this paper another approach to simulating the time series of the average daily BISCR is proposed.

The model of the time series $\vec{B}$ is constructed in the framework of the well-known method of inverse distribution functions (see, for example, Ogorodnikov and Prigarin 1996). To apply this method, it is necessary to define one-dimensional distributions of B_i, $i = \overline{1,365}$ and the correlation matrix R of the time series $\vec{B}$.

The numerical experiments show that the mixtures

$$g_i(x) = \theta_i \frac{1}{B(\alpha_{1i}, \beta_{1i})} x^{\alpha_{1i}-1}(1-x)^{\beta_{1i}-1} + \\ + (1-\theta_i)\frac{1}{B(\alpha_{2i}, \beta_{2i})} x^{\alpha_{2i}-1}(1-x)^{\beta_{2i}-1},$$

$$0 \leq \theta_i \leq 1, \ i = \overline{1,365}$$

of 2 beta-distributions closely approximate sample distribution densities $s_i(x)$. Thr parameters $\theta_i, \alpha_{1i}, \beta_{1i}, \alpha_{2i}, \beta_{2i}$ were chosen using the maximum likelihood method.

For the simulation of the time-series, the sample correlation matrix R was used.

Let us recall the idea of the method of inverse distribution functions. A random sequence with the given distribution densities $g_i(x)$, $i = \overline{1,365}$ and the given correlation matrix R is simulated in 3 steps. The first step is the calculation of the matrix R' that is a correlation matrix of an auxiliary standard Gaussian process $\vec{B}'$. The entry $r'(i,j)$ $i, j = \overline{1,365}$ of the matrix R' is the solution of the equation

$$r(i,j) = \\ = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G_i^{-1}(\Phi(x)) G_j^{-1}(\Phi(y)) \varphi(x, y, r'(i,j)) dxdy,$$

where $r(i,j)$ is an entry of the matrix R corresponding to $r'(i,j)$, the function $\varphi(x, y, r'(i,j))$ is the distribution density of a bivariate Gaussian vector with zero mean, variance equal to 1 and the correlation coefficient $r'(i,j)$ between components number i and j, $\Phi(\cdot)$ is the CDF of a standard normal distribution, G_i is the CDF corresponding to the distribution density $g_i(x)$ of air temperature and atmospheric pressure. Then a standard Gaussian sequence $\vec{B}'$ with the correlation matrix R' is simulated. Finally, $\vec{B}'$ is transformed to $\vec{B}$:

$$B_i = G_i^{-1}\left(\Phi\left(B_i'\right)\right), \ i = \overline{1,365}.$$

The last two steps are repeated as many times as many trajectories are required.

If the matrix R', obtained at the first step, is not positive definite, it must be regularized. In this paper, a method of regularization based on the substitution of negative eigenvalues of the matrix R' with small positive numbers was used (Ogorodnikov and Prigarin 1996). In most cases, after the regularization done the matrix R' requires normalization. The simulation of the standard Gaussian sequence $\vec{B}'$ was done using the spectral decomposition of the matrix R'.

VALIDATION OF THE MODEL

From now on, σ is a standard deviation of the characteristic under consideration when estimating with real data; RD and SD denote the estimations based on real and simulated data, respectively.

All the characteristics of the time series of the real BISCR were estimated based on meteorological observations data collected in 1966-2015 at weather stations, situated in

different Russian cities, for example, in Tomsk (West Siberia; subarctic cyclonic climate), Sochi (The Black Sea region; subtropical climate), Pogranichniy (Russian Far East; arctic climate).

It should be noted that it is possible to simulate as many trajectories as needed to provide a required accuracy of a characteristic estimation. In this paper, for all estimations based on the simulated data we have attained the accuracy above 10^{-4}, so in all the tables presented, the estimations based on the simulated trajectories are given with significant digits only.

Regularization and normalization of the matrix R' alter the correlation structure of the process $\vec{B}'$ and, therefore, the correlation structure of the simulated joint time series $\vec{B}$. This means that the correlation structure of the simulated process differs from the correlation structure of the real one. It should be noted that this difference is rather small, which, for example, is illustrated in Fig. 1. The difference is much less than the statistical error of estimatimating a small sample of real data.

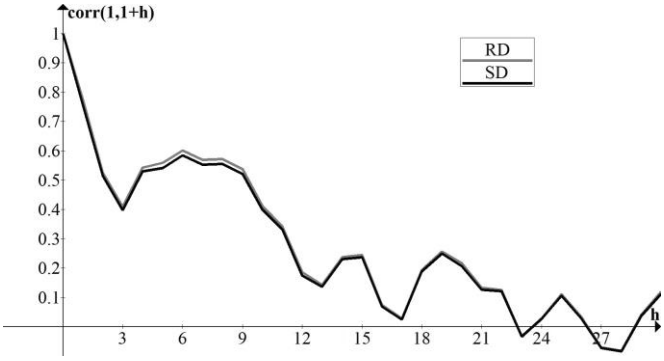


Figure 1: Correlation coefficients $corr(B_1, B_{1+h})$ estimated on real and simulated data. Tomsk

Any stochastic model has to be verified before one starts to use simulated trajectories to study properties of a simulated process. For the model verification, it is necessary to compare the simulated and real data based on estimations of such characteristics, which, on the one hand, are reliably estimated by real data, and on the other hand are not input parameters of the model. Here are several examples of such characteristics.

The first characteristic that was considered was the average number AN of days in a year when the average daily BISCR belongs to the given interval $[a;b]$. Tab. 1 and Tab. 2 show the corresponding estimations. For all of the considered weather stations, simulated data based estimations belong to the corresponding intervals $(AN-2\sigma, AN+2\sigma)$ (in most cases, AN estimated on simulated data belong to the intervals $(AN-\sigma, AN+\sigma)$), so this characteristic is well reproduced with the model proposed.

The other characteristics that were considered were the average monthly minimum MI and the average monthly maximum MA of the BISCR. Fig. 2 shows the corresponding estimations. The numerical analysis shows that for all the considered weather stations and months deviations of estimations on the simulated data from the corresponding estimations based on real data do not exceed σ.

The real and simulated data based estimations of the portion $s(v,n)$ of n-days long time-intervals with the BISCR not exceeding the given level v among all n-days long time-intervals were also compared. Tab. 3 and Tab. 4 show examples of estimations of the portion $s(v,n)$. The estimations based on real and simulated data agree reasonably well.

The results of the numerical analysis show that the trajectories of the model proposed are close in their statistical properties to the real time series of the average daily bioclimatic index of severity of climatic regime at all considered weather stations situated in different climatic zones.

Table 1: Average Number AN of days in a year when the average daily BISCR belongs to the given interval $[a;b]$. Tomsk

$[a;b]$	RD, $AN \pm \sigma$	SD
$[8;10]$	52.240 ± 2.042	53.197
$[7;8)$	99.360 ± 1.750	97.708
$[6;7)$	87.700 ± 1.963	88.122
$[5;6)$	97.400 ± 1.698	97.081
$[4;5)$	27.300 ± 1.919	27.886
$[0;4]$	1.000 ± 0.248	1.005

Table 2: Average Number AN of days in a year when the average daily BISCR belongs to the given interval $[a;b]$. Sochi

$[a;b]$	RD, $AN \pm \sigma$	SD
$[8;10]$	105.440 ± 3.371	105.737
$[7;8)$	193.800 ± 2.918	193.401
$[6;7)$	64.000 ± 1.760	63.401
$[5;6)$	1.740 ± 0.315	2.407
$[4;5)$	0.020 ± 0.033	0.053
$[0;4]$	0.000 ± 0.003	0.000

INFLUENCE OF A CLIMATIC CHANGE

In recent decades there arise many discussions about a climate change and its influence on the human, flora and fauna resources. To construct proper heat-/cold waves prediction systems and to provide a long-range prognosis of the climate regime alteration, it is necessary to study the dependence of different bioclimatic indices on variations of meteorological parameters. In this section, a few results related to studying the influence of a climate change on the real time series of the average daily BISCR and to studying

the sensitivity of the stochastic model to variations of the input parameters are given.

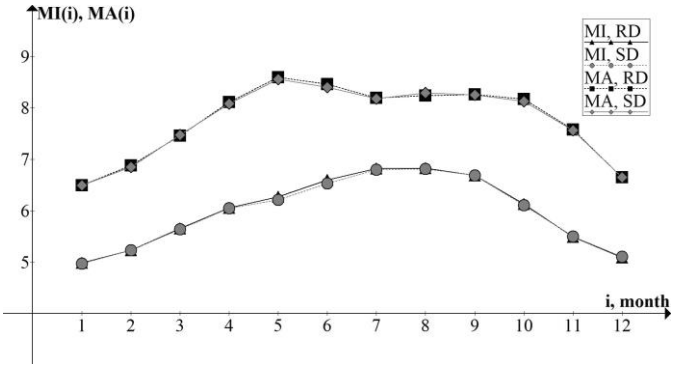


Figure 2: Average Monthly Minimum MI and Maximum MA. Pogranichniy

Table 3: Estimations of $s(v,n)$. Tomsk

n	$v=8$		$v=6$	
	RD, $s(v,n)\pm 3\sigma$	SD	RD, $s(v,n)\pm 3\sigma$	SD
2	0.803 ± 0.020	0.797	0.299 ± 0.020	0.300
4	0.733 ± 0.023	0.722	0.250 ± 0.023	0.251
6	0.685 ± 0.024	0.673	0.217 ± 0.024	0.218
8	0.649 ± 0.025	0.638	0.190 ± 0.026	0.194
10	0.621 ± 0.025	0.610	0.168 ± 0.027	0.174
12	0.598 ± 0.025	0.588	0.149 ± 0.027	0.157
14	0.579 ± 0.026	0.569	0.133 ± 0.027	0.141

Table 4: Estimations of $s(v,n)$. Pogranichniy

n	$v=8$		$v=7$	
	RD, $s(v,n)\pm 3\sigma$	SD	RD, $s(v,n)\pm 3\sigma$	SD
2	0.888 ± 0.021	0.887	0.412 ± 0.017	0.411
4	0.821 ± 0.030	0.818	0.338 ± 0.015	0.338
6	0.769 ± 0.036	0.763	0.305 ± 0.015	0.304
8	0.727 ± 0.039	0.719	0.283 ± 0.016	0.281
10	0.692 ± 0.042	0.682	0.266 ± 0.016	0.264
12	0.663 ± 0.044	0.650	0.251 ± 0.017	0.249
14	0.636 ± 0.044	0.622	0.236 ± 0.017	0.235

Let us study whether the properties of the BISCR time series have changed from 1966 to 2015. All the data collected at weather stations during this period are divided into two independent samples: to the first sample (denoted as RD-1) belong data collected from 1966 to 1990 and to the second one (RD-2) – the data from 1991 to 2015. Tab. 5 and Tab. 6 show estimations of the mathematical expectation μ of BISCR and its variance σ^2 obtained with the samples RD-1 and RD-2. The numerical analysis shows that differences between the values of μ estimated on RD-1 and RD-2 are statistically insignificant. In constrast, the variance σ^2 for most of the considered weather stations and in, around, 85% of days per year have significantly changed. It is not possible to compare estimations of more complex characteristics, obtained on RD-1 and RD-2, because the size of the samples is too small to provide statistically reliable estimations. However, since the stochastic model has already been verified, it is possible to use the simulated trajectories in order to conduct the required study.

Table 5: Mathematical expectation μ of the BISCR. Tomsk

day	RD-1, $\mu\pm 3\sigma$	RD-2, $\mu\pm 3\sigma$
January, 1	5.132 ± 0.280	5.125 ± 0.358
March, 1	5.623 ± 0.363	5.861 ± 0.278
May, 1	7.293 ± 0.521	7.283 ± 0.367
July, 1	7.893 ± 0.426	7.859 ± 0.300
September, 1	7.235 ± 0.246	7.433 ± 0.331
November, 1	5.985 ± 0.337	6.183 ± 0.226

Table 6: Variance σ^2 of the BISCR. Tomsk

Day	RD-1, $\sigma^2\pm 3\sigma$	RD-2, $\sigma^2\pm 3\sigma$
January, 1	0.218 ± 0.142	0.381 ± 0.204
March, 1	0.362 ± 0.268	0.214 ± 0.171
May, 1	0.751 ± 0.280	0.388 ± 0.245
July, 1	0.518 ± 0.288	0.272 ± 0.196
September, 1	0.185 ± 0.177	0.320 ± 0.338
November, 1	0.326 ± 0.345	0.149 ± 0.142

The main idea of the numerical experiment is rather simple. The samples RD-1 and RD-2 are used to define the input parameters of the model and the simulated trajectories are used to study the properties of the time series. Tab. 7 and Tab. 8 show the average monthly minimum MI and the average monthly maximum MA of the BISCR, respectively. As before, all the estimations are given with the significant digits only. It could be clearly seen that there is a slight increase in the values MI and MA. Similar results are also observed at other meteorological stations. This means that in previous 50 years there was a tendency to improve the climatic regime on the territory of the Russian Federation. This also means that the model, presented in this paper, is rather sensitive to small variations of its input parameters.

Table 7: Average monthly minimum MI. Tomsk

Month	RD-1	RD-2
January	4.216	4.353
March	5.084	5.272
May	5.993	6.400
July	6.835	6.933
September	6.150	6.230
November	4.775	4.876

CONCLUSION

In general, the model proposed reproduces quite well the properties of the real time series of the average daily bioclimatic index of severity of climatic regime.

In the future, it is planned to make the model proposed fully parametric. The usage of an analytical approximation of the

sample correlation matrix instead of the sample correlation matrix as it is along with climate change scenarios, will let us carry out more complex numerical experimets related to the study of the influence of a climatic change on the BISCR. It is also planned to compare the model, considered in this paper, with the stochastic model based on the simulation of the joint non-stationary time series of air temperature and relative humidity, wind speed modulus and atmospheric pressure.

Table 8: Average monthly maximum MA. Tomsk

Month	RD-1	RD-2
January	5.947	6.116
March	7.196	7.439
May	8.796	9.042
July	8.787	8.748
September	8.199	8.295
November	6.559	6.654

ACKNOWLEDGEMENTS

This work was partly financially supported by the Russian Foundation for Basic Research (grant No 18-01-00149-a), by the Russian Foundation for Basic Research and the government of the Novosibirsk region according to the research project No 19-41-543001-r_mol_a.

REFERENCES

Akenteva M. S., N. A. Kargapolova and V. A. Ogorodnikov. 2019. "Numerical study of the bioclimatic index of severity of climatic regime based on the stochastic model of the joint meteorological time series". In *Proceedings of the Int. Workshop "Applied Methods of Statistical Analysis" (Novosibirsk Septenber 18-20, 2019)* (in print)

Anderson G. B., M. L. Bell and R. D. Peng. 2013. "Methods to Calculate the Heat Index as an Exposure Metric in Environmental Health Research". In *Environmental Health Perspectives*, Vol. 121, No 10, 1111-1119.

Belkin V. S., M. B. Dyurgerov, A. F. Finaev and S. I. Soroko. 2016. "Bioclimatic evaluation of the human discomfort level for several Antarctic regions". In *Human Physiology*, Vol 42, No 2, 119-127.

Belkin V. S. and G. I. Poltorak. 1983. "Some biomedical aspects of studying mountain regions of Tajikistan". In *Congress of the Geographical Society in Dushanbe*, 19-21. (in Russian)

Belkin V. S., L. N. Sokolov and A. F. Finaev. 1989. "Special features of the bioclimatic zoning of mountain areas". In *Abstracts of the All-Union meeting on human bioclimatology*, 15-16. (in Russian)

Finaev A. F. 2004. "Assessment of the impact of climatic conditions on human labor". In *Archeology and Paleoecology of Eurasia* / Ed. by A.P. Derevyanko. SB RAS, Novosibirsk, 359-376. (in Russian)

Gosling S.N., G. R. McGregor and J. A. Lowe. 2009. "Climate change and heat-related mortality in six cities. Part 2: climate model evaluation and projected impacts from changes in the mean and variability of temperature with climate change". In *Int J Biometeorol*., Vol.53, No 1, 31-51.

Hill L. E., T. C. Angus and E. M. Newbold. 1958. "Further experimental observations to determine the relations between kata cooling powers and atmospheric conditions". In *J. Ind. Hyg.*, Vol. 10, 391-407.

Kershaw S. E. and A. A. Millward. 2012. "A spatio-temporal index for heat vulnerability assessment". In *Environ. Monit. Assess.*, Vol. 184, 7329-7342.

Ohashi Y., Y. Kikegawa, T. Ihara and N. Sugiyama. 2014. "Numerical Simulations of Outdoor Heat Stress Index and Heat Disorder Risk in the 23 Wards of Tokyo". In *J. Appl. Meteor. Climatol.*, Vol. 53, 583-597.

Ogorodnikov V. A. and S. M. Prigarin. 1996. *Numerical Modelling of Random Processes and Fields: Algorithms and Applications.* VSP, Utrecht.

Osczevski R. and M. Bluestein. 2005. "The New Wind Chill Equivalent Temperature Chart" In *Bulletin of the American Meteorological Society*, Vol. 86, 1453-1458.

Revich B. and D.A. Shaposhnikov. 2016. "Cold waves in the southern cities of the European part of Russia and premature mortality of the population". In *Problems of forecasting*, No 2, 125-131. (in Russian)

Shartova N., D. Shaposhnikov, P. Konstantinov and B. Revich. 2018. "Cardiovascular mortality during heat waves in temperate climate: an association with bioclimatic indices". In *Int. J. of Environmental Health Research,* Vol. 28, No. 5, 522-534.

Steadman R. G. 1994. "Norms of apparent temperature in Australia". In *Aust. Met. Mag*, V. 43, 1-16.

EVALUATION OF ALGORITHMS FOR FORECASTING OF INSECT POPULATIONS

Matthias Becker
FG Human Computer Interaction
Gottfried Wilhelm Leibniz University Hannover
Appelstr. 9A, D-30167 Hannover, Germany
E-mail: xmb@hci.uni-hannover.de

KEYWORDS
System Analysis, Stochastic, Decision Support System, Model Evaluation, Agriculture.

ABSTRACT

Pests in Greenhouses can cause major economic damage. A modern method of minimizing this damage is integrated pest management. In this process various chemical and biological pest control methods are combined and their use is supported by a decision support system. These decision support systems are usually based on simulations of the future development of pests and the effects of control measures. A frequently used model for the simulation of insect populations is the escalator boxcar train model, a continuous simulation model based on partial differential equations.
In this work, discrete population-based simulation models have been developed as an alternative. A time-discrete and a discrete event simulation model have been implemented for the simulation of the population development of an insect species and for the predator/prey interaction of two insect species. The population development of the greenhouse whitefly Trialeurodes vaporariorum has been simulated with these simulation models and validated with data from greenhouse experiments. The simulated population developments for the first four weeks are close enough to the experimental data, so that the simulations can be used in the context of a decision support system for the grower.
Based on the validated models for the single species, the interaction of the greenhouse whitefly with the parasitic wasp Encarsia formosa has been modeled and simulated. It has been shown that with a suitable selection of parameters the actual population trend can be approximated, so that the algorithms can be used in a decision support system for pro-active organic pest control.

INTRODUCTION

In modern plant production systems, pest control is a crucial aspect for the economic success of the grower. Pest control means monitoring pest populations in an ecosystem and taking countermeasures to limit the pest population if necessary (Schläpfer 1999, Wilby and Thomas 2002). Several methods for the control of the population of herbivorous insects are available, namely, chemical pest control, natural pest control, and biological pest control (Wilson and Tisdell 2001, Price and Martinsen 1994). The chemical pest control in the agricultural field is wide-spread and consist of the use of pesticides treatments being applied in the greenhouse or in the environment (Geiger et al. 2010, Marrone 2009). In the long run, the targeted herbivorous insects could develop resistance towards the pesticides used. Gradually, more harmful pesticides are developed to efficiently kill off the herbivorous insects. Wilson and Tisdell show that continuous pesticides treatment may lead to distortion of agricultural systems (Wilson and Tisdell 2001). Thus other methods are needed, and especially in organic farming one of the goals is to limit chemicals and rely on alternative methods for natural and biological pest control. Natural pest control is the control of parameters such as (Jones 2015) fungus, illumination, wind, temperature, wave and etc., in order to make the environmental parameters not suitable for pest development.
Biological pest control comprises the introduction of predators, parasites, and/or pathogens in order to control the population of herbivorous insects in a certain environment (Marrone 2009). Predators are insects that actively feed on various species of insects as source of nutrients. Parasites are insects that reproduce by laying their eggs on/in the body of other insects. As a result, the larva grow by feeding on the host. The predator or parasitic insect species that are used in biological pest control are called beneficial insects or beneficials for short. The methods used to increase the population of beneficial insects used in biological pest control can be sub-categorized into three methods, namely, augmentation, conservation, and importation (Jones 2015). Biological pest control is an important component in organic farming, since in organic farming introduction of most chemical pesticides are not allowed.

Decision Support System in Agriculture

Decision support systems for pest management have a long history, e.g.PMEX (Pest Management Executive System), PETE (Predictive Extension Timing Estimator), and BIOSCHED (Pest Biological Scheduling System) (Edens and Klonsky 1977, Welch et al. 1978, Gage et al. 1982) and DSS for fruit production in Pennsylvania (Hull et al. 1992, Rajotte et al. 1992). Currently,

SOPRA is the most sustainable DSS implemented for fruit production in Switzerland and southern Germany (Höhn et al. 2007, Samietz et al. 2015; 2011). A crucial component in the DSS is the algorithm for simulating and forecasting the development of the currently observed pest population. Several algorithms for that purpose exist, with various accuracy and demands for computing time and memory usage.

SIMULATION ALGORITHMS FOR INSECT POPULATIONS

The modeling of insect populations is an important research topic and many approaches exist. Main factors that influence the population dynamics of insects are temperature, length of daylight, and light intensity (Shimoda and Honda 2013, Wyatt and Brown 1977). Wyatt and Brown observed that *Aphis gossypii* reproduces at a higher rate in low temperature. Hence, these factors must be taken into account in order to obtain an accurate result of insect population simulation. In order to accommodate multiple factors, Hemerik and van Nes have proposed MATLAB insect simulation (INSIM) (Hemerik and van Nes 2008; 2016) based on boxcar train. Besides that, the analysis (Wubs et al. 2014) have mentioned that the INSIM simulation could only provides similar trends compared to the actual population development monitored on the yellow sticky traps. In Yeow and Becker (2018) the use of these models for pro-active pest control has been studied. The aforementioned simulation model requires the temperature dependent development rates, α, relative mortality rates, β, and relative birth rates, η for each stage of the pest to generate the population prognosis.Concerning the simulation algorithms, the community basically distinguishes between continuous and discrete models. Continuous simulation models rely on the description of the model's behaviour based on differential equations. Differential equations describe a continuous model development, which however has to be discretized for analysis of the model. The discretization of the continuous models might introduce a long runtime if the time step is chosen to small, or might result in inaccurate results if the timestep is chosen to large. To circumvent those problems, a model can be formulated as discrete event model, where only the points in time are considered when a change of the model state happens. The simulations execution then only hops from event to event, computing the new state of the simulation only when actually something is happening.

Population-based Discrete-Event Simulation

The boxcar train model for a single species has been implemented in Yeow and Becker (2018) already. As next step we build a discrete-event simulation for the same case. The details can be found Wahlmann (2018), where also the following figures origin. The crucial parameters here are the same as in the continuous simulations: the number of developmental stages of the species, the

```
void simulateSingleStep(int step, double stepsize, double minTemp, double maxTemp
    , List preyPopulation, List predatorPopulation, List interactable) {
  double currentTemp = ((maxTemp + minTemp) / 2) - (((maxTemp - minTemp) / 2) *
      cos(2 * PI * stepSize * step));
  simulatePrey(preyPopulation, stepSize, currentTemp);
  simulatePredator(predatorPopulation, interactable, stepSize, currentTemp);
}
```

Figure 1: One time step of continuous predator/prey simulation

temperature-dependent development and mortality rate of each stage, as well as the female ratio and egg-laying frequency of the adult stage.
In the population-based discrete-event Simulation we keep track of each single individuum and its state. For each simulation iteration, the point in time of the next change of state of each individual's development is calculated.

Predator/Prey interaction
We extend the model by a beneficial insect species which acts as predator and uses the pest as prey. In this work we concentrate on white fly and Encarsia formosa, as mentioned above.

Continuous Model
For the continuous simulation the model had to be extended by two lists, namely the population of the predator, second the individuals of the prey species, that are amenable to interaction with the predator. For initialisation, the prey species is initialized with a population spread over all stages (as this reflects reality). The predator population is assumed to be inserted at a certain point in time in adults stage, as this is the practice in reality. One simulation step consists of the measures seen in fig. 1.
The simulation of the prey species runs as in the single species case, however extended by a binary variable that indicates, whether an individual is parasitized or not. If a prey insect dies due to the mortality rate, it is checked whether it is parasitized, if yes then also the parasite dies. Furthermore parasitized larvae do not continue their development: they stay in the current stage.
The simulation of the predator follows fig. 2. The development of the larvae happens completely inside the prey species. We assume that the prey dies if the parasite dies and vice versa. Note that E. formosa does only develop further when the white fly has reached L4 stage. When modelling the hatching of an adult predator, the prey has to be removed from the simulation. For the reproduction of the predator we assume a certain number of eggs carried by the adult (max. 8,9) and laid by the rate found in literature van Roermund (1995). The most prominent action of E. formosa is the interaction of the adult encarsia with the larvae of the white fly. Assuming a Poisson distribution we calculate the probability, that the will be an interaction, as described in van Roermund (1995). Additionally it is assumed that the interaction needs a certain time and the next interaction happens strictly afterwards. If on that base it

```
void simulatePredator(List pedatorPopulation, List interactable, double stepSize,
     double temp) {
List newPredatorPopulation = new List();
for (Insect insect : predatorPopulation) {
boolean died = false;
if (!(insect.getStage().equals(predatorReproducingStage)) {
advanceDevelopment(insect, temp, stepSize);
if (insect.getDevelopment > 1) {
died = checkMortality(insect, temp);
if (died) {
interactable.remove(insect.getHost());
preyPopulation.remove(insect.getHost());
}
advanveStage(insect, temp);
}
} else {
insect.updateEggload(temp, stepSize);
double interactionProbability = 1 - exp(-1 * interactionRate * interactable.size
    () * stepSize);
if (insect.getInteractionTime() < 0); {
if (random.nextDouble() < interactionProbability) {
int preyId = random.nextInt(interactable.size());
interact(insect, interactable.get(preyId), stepSize, newPredatorPopulation);
}
} else {
insect.setInteractionTime(insect.getInteactionTime() - stepSize);
}
if (insect.getDevelopment > 1) {
died = true;
}
}
if (!died) {
newPredatorPopulation.add(insect);
}
}
predatorPopulation = newPredatorPopulation;

}
```

Figure 2: Predator simulation

is decided that an interaction happens one prey individual will be chosen randomly from the corresponding list of suitable individuals in the right stage. Moreover, there is no multiple parasitizing and no host feeding on parasitized larvae. See fig. 3

Discrete-Event Model

In the discrete-event simulation, the same mechanisms have to be integrated as mentioned in the section before. Again, a list of parasitizable individuals is added and the events have been adapted, obeying the mechanisms mentioned above concerning the non-development of parasitized larvae, mutual deaths of prey/predator etc. The calculation of the interaction and interaction probability is slightly different compared to the continuous simulation, since we have to integrate over the time between events, for details see fig. 4.
The necessary biological mechanisms as well as the quantitative values have been taken from literature: the development parameters of the white fly from Wubs et al. (2014), the development parameters of Encarsia Formosa from van Roermund (1995), and the probabilities for laying eggs into different larval stages of white fly from Nechols and Tauber (1977).

RESULTS

After thorough verification of our implementation we conducted simulations with input data from real experiments and validated our simulation models by compar-

```
void interact(Insect predator, Insect prey, double stepsize, List
    newPredatorPopulation) {
if (prey.isParasitized()) { // Antennal rejetion
predator.setInteractionTime(50s);
return;
}
double parasitizationProbability = getParasitizationProbability(prey.getStage());
double hostFeedingProbability = getHostFeedingProbability(prey.getStage());
double r = randon.nextDouble();
if (r < parasitizationProbability && predator.getEggLoad() > 1) { // Oviposition
Insect parasite = new Insect("Encarsia formosa", "White", 0);
parasite.setHost(prey);
prey.setParasite(parasite);
prey.setParasitized(true);
newPredatorPopulation.add(parasite);
predator.setEggLoad(predator.getEggLoad() - 1);
predator.setInteractionTime(360s);
} else if (r > 1 - hostFeedingProbability) { // Host feeding
preyPopulation.remove(prey);
interactable.remove(prey);
predator.setInteractionTime(1600s);
} else { // Ovipositional rejection
predator.setInteractionTime(350s);
}
}
```

Figure 3: predator simulation detail

```
void interact(Insect predator, Insect prey) {
   if (prey.isParasitized()) { // Antennal rejection
      predator.setNextEvent(time + calculateInteractionTime(50s));
      return;
   }
   double parasitizationProbability = getParasitizationProbability(prey.getStage
       ());
   double hostFeedingProbability = getHostFeedingProbability(prey.getStage());
   double r = random.nextDouble();
   if (r < parasitizationProbability && predator.getEggLoad() > 1) { //
       Oviposition
      Insect parasite = new Insect("Encarsia formosa", "White", time +
          calculateLifetime());
      parasite.setHost(prey);
      prey.setParasite(parasite);
      prey.setParasitized(true);
      population.add(parasite);
      predator.setEggLoad(predator.getEggLoad() - 1);
      predator.setNextEvent(time + calculateInteractionTime(360s));
   } else if (r > 1 - hostFeedingProbability) { // Host feeding
      removeHost(prey);
      predator.setNextEvent(time + calculateInteractionTime(1600s));
   } else { // Ovipositional rejection
      predator.setNextEvent(time + calculateInteractionTime(350s));
   }
}
```

Figure 4: Discrete-Event simulation of the interaction

ison with the output of our simulation with the actual development of the populations in the real world experiment in the greenhouse. In the real experiments, first a population of white flies is given the possibility to build up. For the first experiment, 100 E. formosa have been set into the greenhouse when 51 white fly have been counted. The minimum and maximum temperature data has been recorded and been fed into the simulation. the interaction rate have been set to 0.001 and 0.002, for discrete-event and continuous simulation respectively. The results can be observed in fig. 5 from Wahlmann (2018). Note that the population of the E. formosa has been scaled by 1/100. The population of whitefly is still growing, before the deployment of encarsia takes effect (WF Exp., red line). Both simulations reproduce the experimental development, however with some days delay. The discrete-event simulation (green line) shows a smoother population dynamics. The sim-

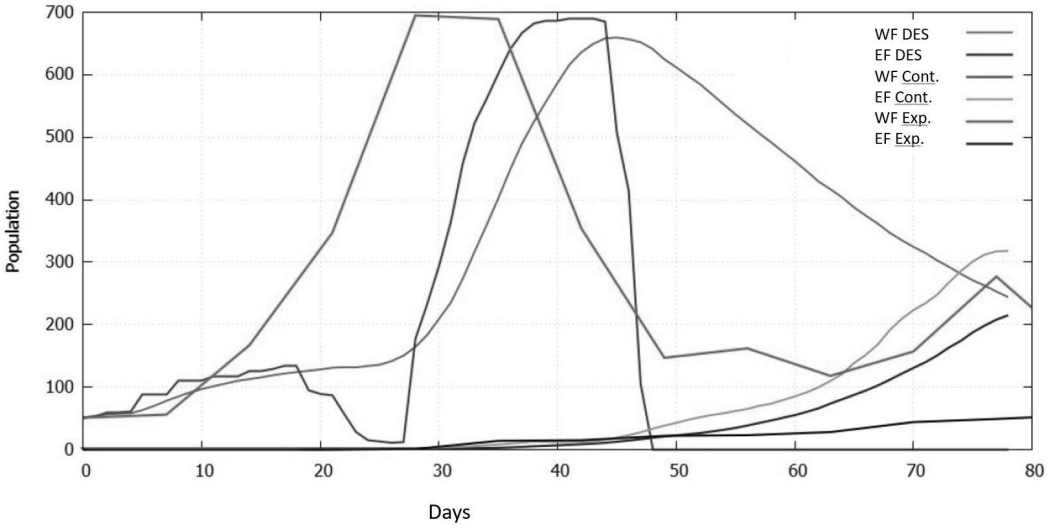


Figure 5: Validation of predator-prey interaction

ulated population of E. formosa also follows the development of reality, however both simulation algorithms overestimate the population of E. formosa. (blue and cyan line).

CONCLUSION

In this work we study different simulation algorithms for modelling and simulation of two interacting populations of insects. The context of this work is a decision support system for maintainers of greenhouses. Based on the currently observed pest count inside a greenhouse, the decision support system estimates the development of the pest population in the near future. If a certain threshold is met, then countermeasures consisting of insertion of a beneficial insect species should be advised. The success of the countermeasures should also be estimated and reported by the decision support system. The application of pesticide does not need to be simulated, since the correct application of pesticides should result in a complete removal of the pest. However in organic farming, pesticides are not allowed and alternate methods have to be use. One of them is the introduction of beneficial insects, that feed on the pest. The deployment of such methods is more sophisticated and the right dose and effect of the dose is harder to estimate. Here our simulation methods come into play: Based on the current pest count, the decision support system can answer the what-if questions: What if no measure is take? What if a certain number of beneficials is introduced? Is the dose enough to stop the pest development or keep it under the economic damage threshold level?. We build two simulation models: one discrete-event model and one model based on differential equations. We observed that both models can reproduce experimental data to a certain extent. The continuous model needs more runtime, however both computational requirements are reasonable low concerning our application case. The accuracy of the simulations is difficult to judge, since a natural system shows more variation and has much more unknown variables and influence on the system. In our model (and also in most models in that area) we do not include possible other species competing for food. Also unforeseeable disturbances are for obvious reasons not included. Examples are: A huge pest population enters from outside through a window. Or additional beneficials find their way into the greenhouse. It can be observed quite often, that e.g. greenhouses in rural areas have many natural invaders making it difficult to make precise predictions using simulations. However our experiments showed that both algorithms have reasonably high accuracy for the application case. Note that differently to technical simulations, where a high accuracy can and needs to be achieved (think of security related domains such as car manufacturing industry), in our application case we need more like an estimate of the pest development, and an estimate of the effect of possibly introduced beneficials. Furthermore, there are also other techniques that could be used in the predictive pest management system, such as, machine learning. The advantage of modeling the population development by machine learning techniques is, that no explicit simulation model has to be constructed and validated out of the collected data; instead, the collected data is directly learned into a black box model and can be used immediately for prediction.

ACKNOWLEDGEMENT

The project is supported by funds of the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the innovation support programme.

REFERENCES

Edens T. and Klonsky K., 1977. *The pest management executive system (PMEX): Description and economic summary. Mich State Univ Agr Exp Sta Rept.*

Gage S.H.; Whalon M.E.; and Miller D.J., 1982. *Pest event scheduling system for biological monitoring and pest management. Environmental Entomology*, 11, no. 6, 1127–1133.

Geiger F.; Bengtsson J.; Berendse F.; Weisser W.W.; Emmerson M.; Morales M.B.; Ceryngier P.; Liira J.; Tscharntke T.; Winqvist C.; et al., 2010. *Persistent negative effects of pesticides on biodiversity and biological control potential on European farmland. Basic and Applied Ecology*, 11, no. 2, 97–105.

Hemerik L. and van Nes E.H., 2008. *A new release of INSIM: A temperature-dependent model for insect development.* In *Proceedings of the Netherlands Entomological Society Meeting.* vol. 19, 147–155.

Hemerik L. and van Nes E.H., 2016. *Predicting the potential establishment of two insect species using the simulation environment INSIM (INsect SIMulation). Entomologia Experimentalis et Applicata*, 159, no. 2, 222–229.

Höhn H.; Höpli H.; Schaub L.; Samietz J.; and Graf B., 2007. *SOPRA: Phenology Modelling of Major Orchard Pests–from Biological Basics to Decision Support.* In *VIII International Symposium on Modelling in Fruit Research and Orchard Management 803.* 35–42.

Hull L.; Eby V.; Heinemann P.; and Crassweller R., 1992. *A working description of the Penn State Apple Orchard Consultant, an expert system. Plant disease*, 76, no. 6, 545.

Jones M.M., 2015. *Biological pest control.*

Marrone P.G., 2009. *Barriers to adoption of biological control agents and biological pesticides. Integrated pest management Cambridge University Press, Cambridge*, 163–178.

Nechols J.R. and Tauber M.J., 1977. *Age-specific interaction between the greenhouse whitefly and Encarsia formosa: influence of host on the parasite's oviposition and development. Environmental Entomology*, 6, no. 1, 143–149.

Price P.W. and Martinsen G.D., 1994. *Biological pest control. Biomass and Bioenergy*, 6, no. 1-2, 93–101.

Rajotte E.; Bowser T.; Travis J.; Crassweller R.; Musser W.; Laughland D.; and Sachs C., 1992. *Implementation and adoption of an agricultural expert system: The Penn State Apple Orchard Consultant.* In *III International Symposium on Computer Modelling in Fruit Research and Orchard Management 313.* 227–232.

Samietz J.; Changins-WÃ H.H.; Changins-WÃ E.R.; Changins-WÃ L.S.; Changins-WÃ B.G.; et al., 2015. *Decision Support for Sustainable Orchard Pest Management with the Swiss Forecasting System SOPRA. Acta horticulturae*, , no. 1099, 383–390.

Samietz J.; Graf B.; Hohn H.; Schaub L.; Hopli H.U.; and Razavi E., 2011. *Web-based decision support for sustainable pest management in fruit orchards: development of the Swiss system SOPRA.* In *Efficient Decision Support Systems-Practice and Challenges From Current to Future*, InTech.

Schläpfer F., 1999. *Expert estimates about effects of biodiversity on ecosystem processes and services. Oikos*, 346–352.

Shimoda M. and Honda K.i., 2013. *Insect reactions to light and its applications to pest management. Applied Entomology and Zoology*, 48, no. 4, 413–421.

van Roermund H., 1995. *Understanding biological control of greenhouse whitefly with the parasitoid Encarsia formosa: from individual behaviour to population dynamics.*

Wahlmann D., 2018. *Verfahren zur Simulation von Insektenpopulationen.* Master's thesis, Leibniz University Hannover, Germany.

Welch S.; Croft B.; Brunner J.; and Michels M., 1978. *PETE: an extension phenology modeling system for management of multi-species pest complex. Environmental entomology*, 7, no. 4, 487–494.

Wilby A. and Thomas M.B., 2002. *Natural enemy diversity and pest control: patterns of pest emergence with agricultural intensification. Ecology Letters*, 5, no. 3, 353–360.

Wilson C. and Tisdell C., 2001. *Why farmers continue to use pesticides despite environmental, health and sustainability costs. Ecological economics*, 39, no. 3, 449–462.

Wubs M.; Böckmann R.; Meyhöfer R.; and Hemerik L., 2014. *Modelling development of the whitefly Triakeurodes vaporariorum for a decision-support-system.* In *Proceedings of the Netherlands Entomological Society meeting.* vol. 25.

Wyatt I. and Brown S.J., 1977. *The influence of light intensity, daylength and temperature on increase rates of four glasshouse aphids. Journal of Applied Ecology*, 391–399.

Yeow K.W. and Becker M., 2018. *JIS: Pest Population Prognosis with Escalator Boxcar Train.* In *2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM).* 381–385.

DEGREE CENTRALITY AND THE PROBABILITY OF AN INFECTIOUS DISEASE OUTBREAK IN TOWNS WITHIN A REGION

Elizabeth Hunter and John Kelleher
School of Computer Science
Technological University Dublin
Kevin Street
Dublin 8
Ireland
email: elizabeth.hunter@mydit.ie

Brian Mac Namee
School of Computer Science
University College Dublin
Belfield
Dublin 4
Ireland

KEYWORDS
Agent-based model, Health Sciences, Discrete Simulation, Compartmental, Centrality

ABSTRACT

Agent-based models can be used to help study the spread of infectious diseases within a population. As no individual town is in isolation, commuting patterns into and out of a town or city are a vital part of understanding the course of an outbreak within a town. Thus the centrality of a town in a network of towns, such as a county or an entire country, should be an important influence on an outbreak. We propose looking at the probability that an outbreak enters a given town in a region and comparing that probability to the centrality of the town. Our results show that as expected there is a relationship between centrality and outbreaks. Specifically, we found that the degree of centrality of a town affected the likelihood of an outbreak within the network spreading to the town. We also found that for towns where an outbreak begins the degree of centrality of the town affects how the outbreak spreads in the network.

INTRODUCTION

The spread of infectious diseases is often thought of as a potential threat to global security. While some believe the biggest threat is of a yet unknown disease there are many well known diseases that are spreading throughout the world today. Diseases previously thought near eradication, such as measles and mumps, have recently re-emerged as a threat due to a number of factors such as lack of vaccinations and global travel. The 2018-2019 influenza season was one of the longest seasons in recent years and the 2019-2020 season is predicted to be worse than normal; at the same time recurring outbreaks of diseases, such as Ebola in Africa, have the risk of spreading abroad; and the lack of sanitary conditions in certain parts of Los Angeles is leading to fears of the spread of diseases such as typhus. One of the best ways to stop an outbreak and prevent a public health emergency is through preparedness. In many cases it may be impossible to test how an intervention will work in a real outbreak. This is where modelling becomes important. Modelling creates a simplified system to study a more complicated real world system and allows us to test intervention strategies and preventative measures.

Within epidemiology research a number of modelling methods are commonly used, including equation based models, such as compartmental SIR models, and computer simulation models, such as agent-based models. In this paper we adopt agent-based models because they capture interactions between factors and emerging patterns that can affect the spread of an infectious disease. There are many factors that can influence the course of an outbreak. Knowing how factors such as population size or vaccination rates lead to different outbreaks between towns can help to lessen an outbreak when it starts or help to focus preventative strategies in the most vulnerable areas. Hunter et al. (2018) showed that other factors can influence the course of an outbreak and two towns that appear similar on paper might have very different outbreaks. While Hunter et al. (2018) study towns in isolation we propose looking at towns within a larger network or region, such as a county. Although understanding how towns differ when their population is closed is important in understanding which towns might be more susceptible to an outbreak if it enters the town, it is also necessary to study how commuting patterns can have an effect on the spread of the outbreak. Knowing how different towns within a network might be more or less susceptible to an outbreak could be essential in helping to stop the spread of a disease when it takes off. Resources might be better funnelled towards towns that have a higher probability of an outbreak in the county then those with a lower chance of an outbreak. We propose that it not just factors within an individual town that might make it more susceptible to an outbreak, such as population density, but the centrality of the town within its network, specifically the transportation network and the number of commuters travelling between each town.

To study this we take an agent-based model for the spread of diseases that is made up of a network of towns. For each town we look at the degree centrality of the town and how that affects the spread of the disease to that town. We first will introduce the model that is used for the study and the different components of the model, then we discuss the experiment: how centrality is calculated and how we account for the other differences in towns. Finally we discuss the results and the conclusions of the study.

MODEL

The model used in the study is a version of the Hunter et al. (2018) model scaled up to model a region made up of many towns instead of a single town. While a town model tells us interesting ideas about the susceptibility of a given town, it is limited in its capacity to capture the town as part of a larger system. In a real world scenario, everyone in the town would not work within the town, instead some would commute out to other towns and others would commute in. This could change the course of the outbreak. To scale up the model we made a number of changes and assumptions that we will outline in this section. We break the description of the model up into the four main components of an agent-based models in human infectious disease epidemiology outlined in Hunter et al. (2017).

Society

Agents are added into the town based on the population data from the Irish Central Statistics Office (CSO) (CSO 2014). The CSO data is at the small area level. Small areas are geographic census areas that contain between 50 to 200 dwellings. They are the smallest area over which the Irish census data is aggregated. For each small area we create a population that reflects the population statistics of that small area including age, sex, household size and economic status. Irish vaccination data is used to determine the percentage of each age group that have received vaccinations for the infectious disease being modelled. For example, if 90% of 1 year olds in Ireland had been given the MMR vaccination in 2011 and we are running a model for 2012, we give each agent in the model with an age of 2 a 90% chance of having been vaccinated. If an agent is vaccinated they are given a 97% chance of being immune to the disease. This takes into account vaccination failure and is based on the vaccine effectiveness rate for measles (Nelson and Williams 2007). Half of the agents with age less than 1 are given immunity to a disease to mimic passive immunity infants receive from their mothers (Nicoara et al. 1999). For any agents that have an age corresponding to a vaccination year not in our data we give a 99% chance of being immune. Prior to vaccination campaigns the majority of the population would have either had or been exposed to childhood diseases, such as measles, leaving them immune in later life. Agents are also given a set of social networks, a family network which connects them to all members of their household, and a school or work network that connects agents to other agents in their school or workplace.

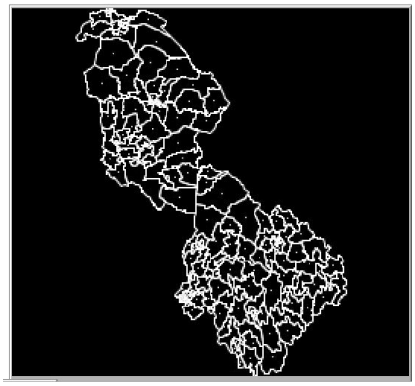


Figure 1: An image of the environmental setup of the model for Leitrim County Ireland

Environment

The counties are created in Netlogo (Wilensky 1999). The Netlogo world is a two dimensional grid where the squares that make up the grid are referred to as patches. Patches in our model represent the small areas within a county. Figure **??** shows the environmental setup of the model for Leitrim County Ireland. All agents that are in a small area at a given time are physically in the same location. However, agents will keep track of their location within that small area. There are four possibilities for agents locations within a small area: home, work, school, or the community. Who an agent comes into contact with depends on that location. For example, if an agent is at home, they know they are at home and will only come into contact with other members of their household who are also at home. Agents in the community within a small area will come into contact with other agents in that small area in the community but will not come into contact with all agents in the community. If two agents from the same household are in the community there is a larger probability of them coming into contact then two agents who are in the same workplace or school who in turn have a larger probability of coming into contact than two agents who have no other connection. All agents within a small area patch have access to information about the patch they are in, including the number of primary and secondary schools in the small area. As well as the real world distances between the center of that small area and all other small areas in the model.

Transportation

Transportation in the town model is a simple model where agents move in a straight line between their current location and their desired destination. When not moving to home, school, or work, agents will pick a random location in the town to move to. It is a naive transportation model but for a small town it is a appropriate simplification. However, we do not feel that this simplification scales to a larger model. Moving randomly throughout a county is not realistic as moving randomly within one town as the distances between locations are much larger. Thus we use a gravity model to determine agent movements. Gravity models are a type of transportation model that is similar in formula to Newton's gravitation model. A traditional gravity model gives the interactions between two location pairs and determines those interactions based on the characteristics of a location and the distance between locations (Rodrigue et al. 2006). In our model, agents will move between home and school or work at certain predetermined times and will return home at predetermined times. On weekends, summers for students and after school or work hours agents will move through the community. An agents movements when they are deemed to be within the community are determined by our gravity model. The probability of an agent moving to another small area is proportional to the population density of the small area, an area that has a lot of other agents is more attractive, and inversely proportional to the distance to the small area from the agents current location, areas that are farther away are less attractive. We feel that this transportation model provides a more accurate simplification of movement within a larger area than that in the original town model.
Commuting patterns within the model are determined using CSO Place of Work, School or College - Census of Anonymity Records (POWSCAR) data (CSO 2017). This dataset provides information on the commuting patterns of people in Ireland and gives the number of people that commute from one electoral division to another. Electoral divisions are the census geographic area one step above the small areas.

Disease

The disease part of the model is the same as that in Hunter et al. (2018) and is based off of an SEIR compartment model that tracks agents' movements between susceptible, exposed, infected and recovered categories. Disease transmissions occurs as follows: if a susceptible agent comes into contact with an infectious agent, the susceptible agent has a percentage chance of becoming exposed. That percentage is used to determine if an agent will move from the susceptible state to the exposed state. This decision is made by drawing a random number between 0 and 1 from a uniform distribution

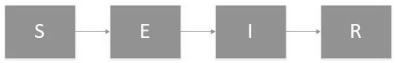


Figure 2: An image representing the disease model. Agents move from susceptible to infected to exposed than recovered

and comparing the random number with the percentage chance of the agent becoming exposed. If the random number drawn is less than the percentage chance of becoming exposed, the agent will move to the exposed state. If the number is greater than the percentage, the agent will remain susceptible. The agent will be in the exposed state for a predetermined period of time and once that period of time is over will move to the infected state. The agent will be in the infected state for a predetermined period of time and then will recover. The agent will stay in the recovered state for the rest of the model runtime and can no longer be infected. A representation of the disease model can be found in Figure **??**.

We do not consider agent death in the model, however, as it is a short time frame, with only one outbreak, we do not feel that agents deaths or births would have a significant impact on the model. In addition, once an agent recovers they are removed from the susceptible, exposed, and infected cycle of the model and thus do not have an impact on the disease dynamics any more.

In this work we study a measles-like disease and the disease dynamics reflect that. On average an individual will stay exposed to measles before becoming infectious for 10 days (Nelson and Williams 2007). The time an agent remains exposed is determined for each agent from a normal distribution with a mean of 10 and a standard deviation of 0.5. Once infectious an individual will remain infectious for an average of 8 days (Nelson and Williams 2007). The time an agent remains infectious in the model reflects this and is determined for each agent from a normal distribution with a mean of 8 and a standard deviation of 0.5. We determine the infection rate, the percentage chance that a susceptible agent will be infected after contact with an infectious agent using the basic reproductive number R_0 for measles (12-18) (Nelson and Williams 2007). The basic reproductive number is defined as the number of individuals infected by one infectious individual in a completely susceptible population. It is the standard measure of transmission of a disease. The parameter can be broken down into three components, number of contacts per unit time (c), the transmission probability per contact (p), and the duration of the infectiousness (d) (Thomas and Weber 2001). As we can determine the number of contacts per unit time from our model and we know the duration of the infectiousness, we can determine the transmission probability per contact.

Schedule

The schedule for our model remains the same as the Hunter et al. (2018) model. The model is run on discrete time steps and runs from initialization to when no more agents are exposed or infected. Each time step represents two hours in a day, with 12 steps representing one day. During weekdays agents will travel to work or school at designated times and remain there until the time assigned for them to leave. Agents can then move throughout the county until they must return home at a certain time. On weekends, and summers for students, all agents can move throughout the county

EXPERIMENT

The goal of the study is to look how the centrality of a town within a region has an influence on an outbreak within the town. To do this we run our model 300 times for the county of Leitrim, Ireland. For each run the initially infected agent is selected at random from the susceptible agents in the county. We then look at how the outbreak spreads through the county and in particular how many outbreaks occur in each town. We look at a set of 16 towns within Leitrim that are made up of more than one small area and are located entirely in Leitrim. The towns are: Ballinamore, Carrigallen, Cloon, Dromahair, Drumod, Drumkeeran, Drumshanbo, Drumsna, Fenagh, Keshcarrigan, Kinlough, Leitrim, Lurganboy, Manorhamilton, Mohill and Tullaghan. An outbreak is defined using the World Health Organization's definition of a measles outbreak, which is two or more linked cases of measles. Thus for each town in the model we determine that there is an outbreak in a given run in that town if two or more agents from the town are infected. We also run a second experiment where we select the town that the outbreak starts in and look at where it spreads from there. Four different towns are chosen to start the infection in, two with high centrality, one with moderate centrality and one with very low centrality.

Centrality

To determine the centrality of the towns we use a weighted degree centrality. There are a number of different ways to determine the centrality of a point in a network. Some of the most common being betweenness, the number of paths that pass through each point, closeness, the proximity of a point to other points in the network, and degree, the number of links between each point in the network. As we want to look at centrality created by commuting between towns degree centrality was chosen, with commutes between different towns considered a link. Degree centrality can be directed based on if the links are coming into the town or out of the town, for the purpose of this study we do not consider the direction of the link. However, a link between two towns in a county where one agent is commuting between the two towns should not have as much of contribution to the centrality of a town as a link between two towns where twenty agents are commuting between the two towns. To account for this we use a weighted degree centrality from Opsahl et al. (2010) that is calculated using a product of the number of links and the average weight of the links adjusted by a tuning parameter. Equation 1 shows the formula for weighted degree centrality with C_i being the centrality of town i, k_i the number of links into the town, s_i the number of agents commuting into or out of the town and α is the tuning parameter. The tuning parameter is used to determine the strength of the weight and the importance of individual link strength: when the tuning parameter is less than one the centrality measure favours more links into the town. If the total number of commuters is fixed a town with more links will have a higher centrality than a town with fewer links. When the tuning parameter is greater than one the centrality measure favours fewer links into the town. If the number of commuters is fixed a town with fewer links will have a higher centrality compared to a town with more links (Opsahl et al. 2010). For the purpose of this study we use an α less than one and set it at 0.5.

$$C_i = k_i * (\frac{s_i}{k_i})^{\alpha} \tag{1}$$

The degree centrality calculated for each town is shown in Table 1.

Table 1: Centrality by town

Town	Centrality
Ballinamore	160.9
Carrigallen	118.9
Cloone	51.4
Dromahair	86.6
Drumod	96.5
Drumkeeran	90.7
Drumshanbo	208.7
Drumsna	75.1
Fenagh	74.5
Keshcarrigan	87.8
Kinlough	55.8
Leitrim	127.3
Lurganboy	58.0
Manorhamilton	232.8
Mohill	212.2
Tullaghan	17.4

We are interested in examining how the centrality of a town affects the outbreaks within a region. To explore this question we wish to control for other factors (town characteristics) other than centrality which may affect an outbreak. Consequently, a key step in our experiment is to identify pairs of towns that differ in terms of centrality but which are similar with respect to other characteristics. Once we have identified these town pairs we can examine how centrality affects a outbreak by

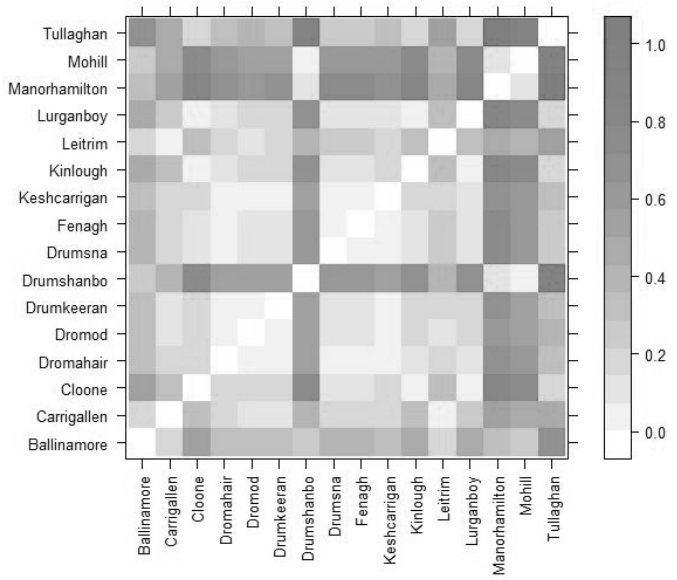


Figure 3: Distance Matrix showing the normalized difference in centrality between two towns

comparing how an outbreak spreads through the region when it starts in different towns. The first step in identifying these town pairs is to identify towns that differ in terms of centrality. To visualise which towns have the biggest difference in centrality we created a distance matrix shown in Figure **??** where the distance is the normalized difference between the centrality for each pair of towns. The more red the square, the greater the difference in centrality and the more white the square the less the difference in centrality between the towns.

Similar Towns

While comparing the results between two towns with different centrality might give an idea of how centrality influences an outbreak within a town, there are many other factors involved. Two towns with markedly different centrality might also be different in size or population density and this could be what is affecting the different results. To control for this we use euclidean distance to find towns that are similar in other characteristics that we believe might influence an outbreak. Each town is represented by a vector of quantitative characteristics: population size, town area (km^2), population density, number of small areas that make up the town, the number of secondary schools, the number of primary schools, the percent of susceptible agents in the town and the percent of agents who are students in the town. All categories except for the number of secondary schools and number of primary schools are standardized. The euclidean distance is then calculated between each of the 16 towns so that we can compare results between similar towns. Figure **??** presents a distance matrix which visualizes which towns are similar based on these categories. The lighter the square the more similar the towns are and the darker the more dissimilar.

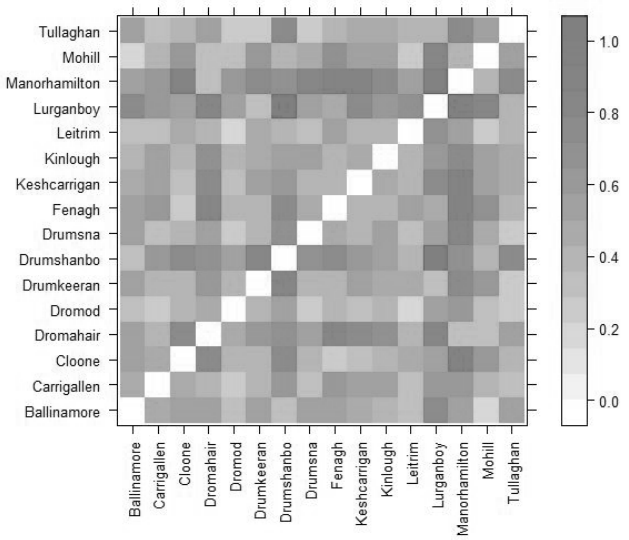


Figure 4: Distance Matrix showing the normalized euclidean distance between towns

To be able to easily focus on the towns that have similar characteristics but different centralities we created a distance matrix that is the difference between the euclidean distance and the difference in centrality, Figure **??** presents this distance matrix. The more blue a square is in the matrix the smaller the euclidean distance and the larger the difference in centrality while the more red a square is the larger the euclidean distance but the more similar the centrality.

RESULTS

After running the model 300 times with a random start location for the outbreak, for each of the sixteen towns we calculate the percent of runs that results in at least one resident becoming infected and the percent of runs that lead to an outbreak (two or more agents infected who are residents of the town). The results along with the centrality for each town can be found in Table 2.
To look at the basic relationship between the degree centrality of each town and the outbreaks that occur in the towns we find the Pearson correlation. The correlation between the centrality of the town and at least one agent from the town getting sick is 0.66 and the correlation between the centrality of the town and an outbreak occurring in the town is 0.65. Although these are not strong correlations, they can be considered to show a moderate relationship (Ratner 2009). This is not unex-

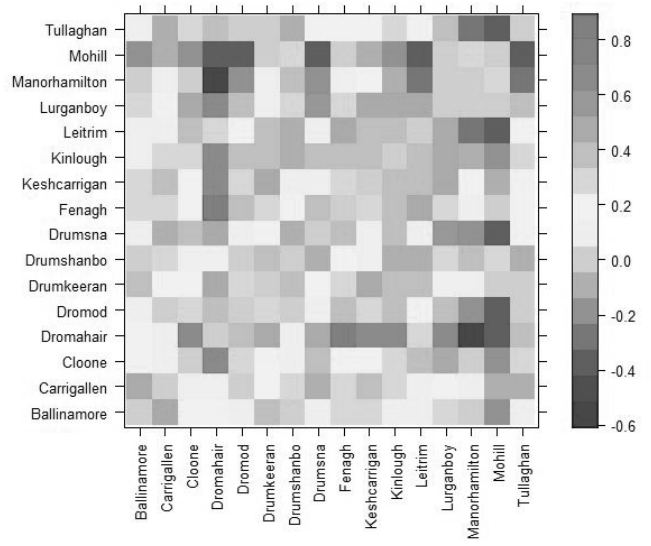


Figure 5: Distance Matrix showing the difference between euclidean distance and centrality between towns

pected as there are other factors besides centrality that influence an outbreak in a town. For example, the number of susceptible individuals in the town or the number of schools and workplaces. From the correlations we can conclude that the degree centrality of the town has a relationship with where the outbreak spreads in the model.

To further look into how the centrality of a town affects the outbreaks within a county we looked at repeated starts of the outbreak in the same town and looked at where the outbreak spreads and compare this to outbreaks that started in a similar town with different centrality. Examining Figure 3 Manorhamilton and Dromahair are similar towns with a large difference in centrality (the square at the intersection of these towns is deep blue). Similarly, Mohill and Tullaghan are similar towns with a large difference in centrality. The centrality of the four towns can be seen in Table 3. Table 4 lists the percent of runs where an infection that started in one of these 4 towns (Tullaghan, Drumahaire, Mohill, and Manorhamilton) resulted in an outbreak occuring in one of the 15 other towns in the Leitrim model. For example, Table 4 shows that 19% of the time that an infection started in Tullaghan and outbreak occurred in Ballinamore.

To get an idea of how the results vary and the outbreak spreads through the network when it starts in specific towns, we compare the outbreaks that occur in the other fifteen other towns where the outbreak did not start with the centralities of those fifteen other towns. Similar to what was done for the initial analysis, we calculate the correlations between the outbreaks for each town in

Table 2: Results by town

Town	Centrality	One Sick	Outbreak
Ballinamore	160.9	38.0	24.0
		(32.5, 43.5)	*(19.2, 28.8)*
Carrigallen	118.9	27.3	17.3
		(22.3, 32.4)	*(13.0, 21.6)*
Cloone	51.4	20.3	7.7
		(15.8, 24.9)	*(4.7, 10.7)*
Dromahair	86.6	32.3	18.3
		(27.0, 37.6)	*(13.9, 22.7)*
Drumod	96.5	31.3	18.3
		(26.1, 36.6)	*(11.3,19.4)*
Drumkeeran	90.7	19.3	7.3
		(15.9,23.8)	*(4.4,10.3)*
Drumshanbo	208.7	34.0	18.6
		(28.6, 39.4)	*(14.3, 23.1)*
Drumsna	75.1	27.7	14.3
		(22.6, 32.7)	*(10.4, 18.3)*
Fenagh	74.5	26.3	11.7
		(21.3, 31.3)	*(8.0, 15.3)*
Keshcarrigan	87.8	22.7	9.3
		(17.9, 27.4)	*(6.0, 12.6)*
Kinlough	55.8	35.3	21.0
		(29.9, 40.7)	*(16.4, 25.6)*
Leitrim	127.3	32.3	18.0
		(27.0, 37.6)	*(13.7, 22.3)*
Lurganboy	58.0	23.7	6.7
		(18.9, 28.5)	*(3.8, 9.5)*
Manorhamilton	232.8	39.7	22.7
		(34.1, 45.2)	*(17.9, 27.4)*
Mohill	212.2	39.0	24.3
		(33.5, 44.5)	*(19.5, 29.2)*
Tullaghan	17.4	30.0	15.3
		(24.8, 35.3)	*(11.3, 19.4)*

Table 3: Correlations between outbreaks started in each town and the centrality of the towns

Town	Centrality
Dromahair	86.6
Manorhamilton	232.8
Mohill	212.2
Tullaghan	17.4

Table 4 and the centralities of these towns from Table 1. The correlations can be found in Table 5 separated by the town the outbreak starts in. The idea being that if the correlations are different we can learn something about how an outbreak spreads through a network based off of the centrality of the the town where it began.

From Table 5 we can see that when the outbreak starts in a town with lower centrality, Dromahair and Tullaghan, there is a moderate correlation between the percent chance of an outbreak in another town and the centrality of that other town. However, the correlations are lower when the outbreak starts in a town with high centrality, Mohill and Manorhamilton. One way to interpret this is that when an outbreak starts in a town

Table 4: Results by town when the outbreak starts in one of four towns

Town	Tullaghan	Dromahair	Mohill	Manorhamilton
Ballinamore	19.0	20.3	24	15.7
	(14.6,23.4)	*(15.8, 24.9)*	*(19.2,28.8)*	*(13.7,22.3)*
Carrigallen	11.0	11.3	16.3	7.7
	(7.5,14.5)	*(7.7,14.9)*	*(12.2,20.5)*	*(4.7,10.7)*
Cloone	8.7	5.3	15.3	5.3
	(5.5, 11.9)	*(2.8,7.9)*	*(11.3,19.4)*	*(2.8,7.9)*
Dromahair	31.3	-	19.3	31
	(26.1, 36.6)	-	*(14.9,23.8)*	*(25.8,36.2)*
Drumod	13.0	15.3	22	7.7
	(9.2, 16.8)	*(11.3,19.4)*	*(17.3,26.7)*	*(4.7,10.7)*
Drumkeeran	10.7	14.7	7.3	15
	(7.2,14.2)	*(10.7,18.7)*	*(4.4,10.3)*	*(11.0, 19.0)*
Drumshanbo	15.3	22.7	18	18.7
	(11.3,19.4)	*(17.9,27.4)*	*(13.7,22.3)*	*(14.3,23.1)*
Drumsna	9.7	13.7	20.3	8
	(6.3,13.0)	*(9.8,17.6)*	*(15.8,24.9)*	*(4.9,11.1)*
Fenagh	8.0	10.3	10.3	7.3
	(4.9,11.1)	*(6.9,13.8)*	*(6.9,13.8)*	*(4.4,10.3)*
Keshcarrigan	8.0	10.7	6.3	6
	(4.9,11.1)	*(7.2,14.2)*	*(3.6,9.1)*	*(3.3,8.7)*
Kinlough	31.3	24.3	17.7	30.7
	(25.8,36.2)	*(19.5,29.2)*	*(13.4,22.0)*	*(25.4,35.9)*
Leitrim	14.3	19.3	18.7	16
	(10.4,18.3)	*(14.8,23.8)*	*(14.3,23.1)*	*(11.9,20.1)*
Lurganboy	24.3	14.3	7.3	27
	(19.2,28.8)	*(10.4,18.3)*	*(4.4,10.3)*	*(22.0,32.0)*
Manorhamilton	40.7	34.3	18	-
	35.1,46.2)	*(29.0,39.7)*	*(13.7,22.3)*	-
Mohill	25.0	19.3	-	16.7
	(19.8,29.5)	*(14,9,23.8)*	-	*(12.4,20.9)*
Tullaghan	-	18	16	26
	-	*(13.7,22.3)*	*(11.9,20.1)*	*(21.0,30.9)*

Table 5: Correlations between outbreaks started in each town and the centrality of the towns

Start of Outbreak	Correlation
Dromahair	0.61
Manorhamilton	0.13
Mohill	0.34
Tullaghan	0.38

with high degree centrality, the centrality of other towns in the network do not have an effect on if the outbreak spreads to that town but when an outbreak starts in a town with low degree centrality the centrality of the other towns have a much higher influence on if the outbreak will spread there. The difference is more pronounced when comparing Dromahair and Manorhamilton versus Tullaghan and Mohill. This could be due to Tullaghan's very low degree centrality. As these are combinations of towns we have identified in the early analysis as having similar characteristics but different centrality we can conclude that most of the difference is due to the centrality and not outside factors.

CONCLUSION

The more that we are able to learn about how outbreaks move throughout a network of connected towns the more prepared we can be for a real world outbreak and not just a simulation. Understanding how well connected a town is in the network of a county or even the entire country might be able to inform the intervention strategies being used. If we know its more likely for an outbreak to spread to the towns with a higher degree centrality interventions could be focused in those areas. Finding out how susceptible a town is to an outbreak could save time and effort when trying to stop an outbreak that has already started or might save costs in focusing vaccination campaigns or other prevention strategies on those towns. Understanding how towns fit into a network of other towns and how that network influences the outbreak in an individual town is important in understanding how infectious diseases spread. Agent-based modelling allows us to understand these differences in towns without having to wait until a real outbreak occurs. In our analysis we were able to find a moderately strong positive linear relationship between the degree centrality of a town and the likelihood that an outbreak would spread to the town. In addition, the analysis shows that where an outbreak starts has an influence on what other towns will be affected. However, further work should be done looking into how other measures of centrality might influence the spread of the outbreak and might interact with degree centrality. For example, does the distance to the town where the outbreak starts affect the results and does closeness centrality interact with degree centrality to influence outbreaks spreading through a network. This could lead to an even greater understanding of how an outbreak will spread through a network and could prove invaluable.

References

CSO, 2014. *Census 2011 Boundary Files*. URL `http://www.cso.ie/en/census/census2011boundaryfiles/`. Date accessed 26.05.2016.

CSO, 2017. *Census 2016 Place of Work, School or College - Census of Anonymised Records (POWSCAR). Central Statistics Office*. URL `https://www.cso.ie/en/census/census2016reports/powscar/`.

Hunter E.; Mac Namee B.; and Kelleher J., 2018. *An open-data-driven agent-based model to simulate infectious disease outbreaks. PLOS ONE*, 13, no. 12, 1–35. 10.1371/journal.pone.0208775. URL `https://doi.org/10.1371/journal.pone.0208775`.

Hunter E.; Mac Namee B.; and Kelleher J.D., 2017. *A Taxonomy for Agent-Based Models in Human Infectious Disease Epidemiology. Journal of Artificial Societies and Social Simulation*, 20, no. 3, 2. ISSN 1460-7425. 10.18564/jasss.3414. URL `http://jasss.soc.surrey.ac.uk/20/3/2.html`.

Nelson K.E. and Williams C.M., 2007. *Infectious Disease Epidemiology Theory and Practice*. Jones and Bartlett Publishers.

Nicoara C.; Zach K.; Trachsel D.; Germann D.; and Matter L., 1999. *Decay of passively Acquired Maternal Antibodies against Measles, Mumps and Rubella Viruses. Clinical and Vaccine Immunology*. .

Opsahl T.; Agneessens F.; and Skvoretz J., 2010. *Node Centrality in Weighted Networks: Generalizing Degree and Shortest Paths. Social Networks*, 245-251, no. 3, 1–35. URL `https://doi.org/10.1016/j.socnet.2010.03.006`.

Ratner B., 2009. *The correlation coefficient: Its values range between +1/1, or do they? Journal of Targeting, Measurement and Analysis for Marketing.* .

Rodrigue J.P.; Comtois C.; and Slack B., 2006. *The Geography of Transport Systems*. Routledge, Taylor and Francis Group, London. ISBN 9780415354400.

Thomas J.C. and Weber D.J., 2001. *Epidemiologic Methods for the Study of Infectious Diseases*, Oxford University Press, chap. Concepts of Transmission and Dynamics. 61–62.

Wilensky U., 1999. *Netlogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University Evanston, IL*. URL `https://ccl.northwestern.edu/netlogo/`.

THE POSSIBILITIES OF SIMULATION IN THE FIELD OF PREPARATION FOR DEALING WITH HYBRID THREATS

MSc. Michaela Jánošíková
Professor Jozef Ristvej, Ph.D.
MSc. Maroš Lacinák, Ph.D.
Dr. Michaela Kollárová
Department of Crisis Management
University of Žilina, www.uniza.sk
Univerzitná 1, 01026 Žilina, SK
E-mail: michaela.janosikova@fbi.uniza.sk

KEYWORDS
Crisis Management, Simulation, Hybrid Threats, Exercises.

ABSTRACT

Hybrid threats have become one of the most discussed security threats of today. In response to increasing risk of such threats in various forms, the European Union has been taking several measures since 2016 to increase the readiness and resilience of systems to face them. The EU has adopted several documents at European and national levels, and has established several institutions dedicated specifically to this issue. In updating its strategic security documents, the Slovak Republic has adopted number of public policies and strategic documents in recent years, that contain a definition of hybrid threats and that set up a new institutional framework for their addressing. Major challenge in this area is preparation of public administration capacities. One of possible tools, that can be used during capacity preparation is simulation. However, such use is only in an early stage of development. In this article, it is described the current stage of the research and implementation of the training of public administration executives through exercises and the exercises themselves. Effectiveness of public policies, documents, institutions and their efforts to tackle hybrid threats will only be tested by practice, but they largely correspond to approach and concept adopted at the EU level.

INTRODUCTION

Safety environment in Europe has undergone major changes in recent years, resulting in a transformation of well-known conventional threats. As a result of technological development, these threats have taken on a whole new dimension and their intensity has increased. Activities, simultaneously threatening the basic attributes of the state or its functionality are referred to as hybrid threats. According to the Framework of the Slovak Republic on Countering Hybrid Threats (2018), the Slovak Republic also faces the spread of disinformation about alleged political and economic problems resulting from the EU and NATO.

A potential hybrid threat against the security of the Slovak Republic lies also in the support of local radical organizations pursuing objectives identical to people, responsible for hybrid threats spreading. One of the main tasks of the public authorities is to evaluate the reported incidents within their competence in relation to hybrid threats, in the light of the indicators, set at the beginning of this chapter. Our aim is to prepare the public administration authorities to address hybrid threats. The urgency of addressing the issue of hybrid threats, that affect the safety environment of the Slovak Republic requires adequate response options to eliminate their manifestations and also effective prevention and preparedness.

HYBRID THREATS AND SIMULATION

In the European Union (EU), considerable attention is payed to the issue of hybrid threats. Since 2016, a number of public policies and documents have been adopted in this area.

Current State of Hybrid Threats

The Joint Framework on Countering Hybrid Threats adopted in 2016 (European, 2016) can be considered as the most important document in this field, that comprehensively addresses the issue and also contains concrete measures and recommendations to mitigate the impact of hybrid threats within the EU. The definition of hybrid threats also appears in this document:

The hybrid threats represent „the mixture of coercive and subversive activity, conventional and unconventional methods (i.e. diplomatic, military, economic, technological), which can be used in a coordinated manner by state or non-state actors to achieve specific objectives while remaining below the threshold of formally declared warfare. There is usually an emphasis on exploiting the vulnerabilities of the target and on generating ambiguity to hinder decision-making processes. Massive disinformation campaigns, using social media to control the political narrative or to radicalise, recruit and direct proxy actors can be vehicles for hybrid threats.“ (European, 2016).

In identifying such threats, we can observe certain specifications. Hybrid threats:

- progress dynamically,

- are created by combination of several factors,
- are coordinated activities of a third party,
- have so called 'snowball effect',
- are below the threshold of usual reaction.

The most common challenges used in the framework of dealing with hybrid threats in context of indicators settings are (Government, 2018):

- external or internal political pressure on top level state officials and state institutions;
- economical or energy pressure as as part of political influence,
- extensive sabotage against key infrastructure,
- cyber-attacks with the potential to cause large scale damage,
- information and propaganda operations to undermine confidence in state institutions, trigger social unrest and severely destabilise political and security stability,
- influencing ethnic, religious and cultural minorities and manipulating them for political purposes,
- threat of using military force.

In Joint Communication to the European Parliament and the Council (2016), the High Representative of the EU for Foreign Affairs and Security Policy and the European Commission identify 22 concrete measures aimed at improving the coordination and readiness of the EU and its individual Member States to identify and address hybrid threats. One of these measures is the Measure no. 19, in which the High Representative together with the European Commission, in coordination with the Member States, will establish a joint operational protocol. They also undertook to carry out regular exercises to improve strategic decision-making capacity in response to complex hybrid threats, in the context of crisis management procedures and integrated policy response to the crisis (European, 2016). Following this recommended measure, an initiative was developed to develop a possible way of preparing public officials to identify and address hybrid threats.

Typology of Simulation and Their Use in the Area of Preparation for Hybrid Threats

According to Rybár et al. (2000), the term simulation can be understood as a process, connected with using already created, prepared and verified model by some subject in dealing with specific task. In terms of crisis management, we are mainly interested in the application of simulation technologies carried crisis phenomena in education, training, practice in all phases of Disaster Management Cycle to crisis events, and research opportunities in this field. Currently, there are three basic types of simulations:

- ***virtual simulation*** means that real people use simulated (virtual) equipment in a simulated environment called a virtual environment,

- ***live simulation*** means that real people use simulated (virtual) equipment in the real world, in particular the various simulators such as aerospace, automobiles, etc. and the use of armed forces for training combat troops,

- ***constructive simulation*** means that simulated (virtual) people use simulated (virtual) equipment in a simulated (virtual) environment (Ristvej et al., 2016).

Virtual simulation is a simulation, based on the use of virtual reality, that refers to interactive computer simulation in a real or imaginary simulated environment. The main advantage of such simulation is that it is able to create a more realistic environment compared to constructive simulations, thereby giving the practitioner a sense of immersion into the given situation (Walle et al., 2010). This type of simulation can be used in the flied of hybrid threats mainly for simulation of cyber-attack.

Live simulation is activity, that involves people performing activities in a quasi-real situation. The advantage of these simulations is that they are more realistic than virtual and constructive simulations. Its main disadvantage is its demandingness in questions of time, finances and also organization. There is also a limited number of possible simulated scenarios per exercise, usually only one scenario can be realized (Walle et al., 2010). The military environment also comes with exercises called Gaming, otherwise referred to as serious games which belong to the category of live simulation. Meesters (2014) states, that the essence of serious games is an action between two or more independent decision-making bodies, seeking to achieve their goals in a predefined context. However, he also states, that not all games are competition between opponents. In some games players cooperate to achieve a common goal, while fighting against an obstacle or situation that arises but is not controlled by another real player. Of course, serious games have much broader context in use, than just the military purposes. This type of simulation can be used in the flied of hybrid threats mainly for simulation of cyber-attack.

Constructive simulation is based on a computer support, where a software tool allows us to simulate processes and phenomena in real or specified time and environment. The aim of creating such an environment is to give the practitioner the impression of performing real operations and activities as if they were performed in the real environment (Simulation, 2019). For its similarity to the military is constructive simulation understood as strategic computer game is often described as a military simulation. Constructive simulation is able to design

and conduct simulated scenarios for most areas dealing with crisis management practically across the whole territory. This type of simulation for teaching and training is called the Computer-aided Exercises (CAX) (Ristvej et al., 2016). Constructive simulation can be used for preparation for hybrid threats in training, where we will prepare scenarios with event tree and we will proceed one selected path.

In addition to these three types of simulation, Rybár et al. (2000) mentions a fourth type: their combination. This type of simulation is commonly used in military trainings. Raiszadeh and Batterson (2018) in their publication mention also the possibility of combining the three types of simulation, referring to such combination as a hybrid simulation.

Hybrid simulation combines constructive, live, and/or virtual simulations, typically in a distributed environment. Such simulations typically combine simulators with actual operational equipment, prototypes of future systems, and realistic representations of operational environments. (Raiszadeh and Batterson, 2018).

One of the possibilities of hybrid simulation is the connection of constructive and virtual simulation. This way it is possible to connect two different simulation environments on the same map background and in the same tactical situation. During the combined exercise, by the constructive simulation, simulated are opposing units and units under control of trained staff, while the crews of individual vehicles can be simulated on virtual simulators. Inputs for trained staff are provided from a unified communication system (Simulation, 2019).

Another option is to connect constructive and live simulation in one exercise. In such exercise, constructive simulation is used to present the situation to the trainees, and may represent a tool to support decision-making within given exercise. The live simulation will be directly related to the plot of constructive simulation and will try to create an environment as close as possible to the real one, as it would be in the case of real phenomena. However, when simulating hybrid threats, individualities of hybrid threats as such must also be taken into account. It should be noted that the hybrid threat is not a sudden situation, but its course is dynamically developing, usually below the threshold of the usual response.

The aim of simulation is to create and proceed scenarios simulating a complex crisis situation, that combines several forms of threats from the category of hybrid threats. These simulations should also include their evaluation and identification of processional procedures that have been effective and successful in dealing with the hybrid threat and also those procedures, that have led to negative results. In the absence of clearly defined protocols and processes on how to proceed in the cases of hybrid threat emergence, the objective is to use hybrid threat simulation to identify the procedures, that are most appropriate to apply in such situations. The aim of the simulation is to identify weaknesses and vulnerability rates within existing competences and tasks of individual public administration institutions (Loveček, 2008).

The current and main task of engaging simulations in the field of preparation for the hybrid threats is their use in the process of preparation of public administration representatives in order to acquire information, received during the theoretical preparation phase and to verify their readiness to implement acquired knowledge into practice.

USE OF SIMULATION IN THE PREPARATION PROCES OF PUBLIC ADMINISTRATION REPRESENTATIVES DEALING WITH HYBRID THREATS

During the preparation of public administration representatives, which is focused on the field of hybrid threats, various types of simulations can be used. These are incorporated into the preparation in order to acquire the information, obtained by representatives during the theoretical preparation phase as well as to verify their readiness to implement the acquired knowledge into practice. However, when selecting the appropriate type of simulation, we must take into account the specific nature of hybrid threats. In particular, hybrid threats have a dynamic course and occur below the usual response threshold. It is therefore necessary to adapt the practical part of the preparation to these characteristics. The practical part of the preparation will be carried out in the form of exercises, supported by constructive and live simulations. Exercises were prepared and will be performed in the Laboratory of Modeling and Simulation of Crisis Phenomena at the University Science Park of the University of Žilina.

Before the exercise, basic theoretical preparation will take place. During this part of the preparation, public officials will obtain basic information about the specific characteristics of hybrid threats, how to identify them, and how to respond and address them.

The main goal of the exercise is to verify the practical application of knowledge about hybrid threats acquired by public administration representatives during training, conducted in the framework of theoretical preparation through a practical dealing with crisis phenomena, as well as to practice crisis communication of practitioners. One of the objectives is also to examine the availability of forces and resources for crisis management, with an emphasis on hybrid threats. In addition to checking the knowledge of practitioners gained from previous theoretical training, a positive aspect is also the reduction of the economic demands of training of public administration officials, compared to the traditional forms of training, according to the experience with using other types of training. The target groups in the prepared exercises are public administration representatives at national and regional level, not local level.

The conception of the exercise consists of several parts, which are fluently, thematically connected. The initiation event will dynamically change during the exercise, the situation will escalate until it will result in an event, with attributes of a hybrid threat. To create an environment, similar to an environment of a hybrid threat, we need to use several forms of simulation of a given event. The exercise will use constructive simulation, through the VR-Forces® software program, a practical exercise with dynamic progress, which will represent live simulation or training of the skills of practitioners to work with verified sources of information. We suppose that the exercise will be completed by theoretical testing, the aim of which is to verify the knowledge of public administration representatives in the solved issues. The results will be evaluated. According to Walker et al. (1989) "Computer-based simulators can bridge the gap between classroom training and live simulations and provide the management training needed".

Constructive simulation will represent an introduction to the situation. Through it, practitioners will see an initiating event, that will gradually, within another tasks grow into a hybrid threat. The event scenario will be simulated through the VR-Forces® software program. Program VR-Forces® is a comprehensive simulation tool that allows us to simulate different situations in real or fictional map background. It contains a vast set of features and models, in attempt to provide as realistic output as possible. The scenario of the initiation event will be based on the leakage of an unknown substance into the water, causing contamination of the drinking water source. Through a constructive simulation, practiced can be the decision-making process of the local crisis staff, as well as communication and coordination of the individual parties and stakeholders, active in resolving the event. As this is a smaller-scale situation, this simulated scenario will only be addressed by local bodies and stakeholders. The signs of the hybrid threat will not manifest until next phases of the scenario.

As the scenario develops, information about water source contamination will start to appear and the media will begin to take advantage of the situation, trying to bring new, not always verified and true information. One of the tasks we have is to verify the ability of trainees to search for the information from relevant databases and their ability to verify them. As a part of the exercise, practitioners will be presented with information in the form of various articles, statements or case studies that may be true or false. Based on information source, message content, etc. it will be for the participants to determine, whether the information is true or false. The accumulation of intentionally generated false information for a specific purpose below the threshold of a normal response also constitutes a kind of hybrid threat. This type of tasks can also be characterized as Serious Game, because the environment, in which the practitioners are located can be changed and adapted to the current situation, created by the decision-making processes of public administration representatives. The environment thus becomes dynamic, which is a characteristic feature of the simulation. The situation will evolve based on the decisions of the trainees, but it will be possible to regulate it through the events that will be continually inserted into the simulation. These will be determined on the basis of a pre-prepared event tree. That should ensure, that the exercise will not move away from the solved problem. Therefore, this practical, dynamic training exercise can be in a way considered a live simulation.

The end of the exercise will be devoted to a theoretical testing, the aim of which will be to verify the knowledge of public administration representatives, obtained by theoretical preparation in the solved field.

Within the exercise, we will investigate whether public administration representatives:

- can identify the presence of hybrid threats,
- have improved their ability to identificaty hybrid threats, as opposed to the pre-training state,
- are able to apply the knowledge of hybrid threats gained during the preparation into practice.

The exercise will also include the evaluation of the exercise as a whole. It will be carried out by staff, responsible for the exercise – Exercise Control Group (ECG) and it will be based on the data, obtained from the exercise evaluation forms. Evaluation of the exercise consists of collecting and processing the information obtained through evaluation reports.

The exercise is evaluated in grades as follows:

- passed (specified exercise tasks were completed in time)
- passed with limitation (most of the specified exercise tasks were completed)
- failed (most of the specified exercise tasks were not fulfilled).

Immediately after the exercise, an evaluation will also be performed in the form of feedback from the trainees. That will be conducted with the use of brainstorming method and questionnaire. Brainstorming will be based on pros and cons of the exercise, while the answers of the trainees will be written down on the board. An important aspect of this form of feedback will be to record all views, whether other participants agree or not. Questionnaire will cover questions about exercise and its use in different areas of public sector. Recorded answers, suggestions and comments will be taken into account in the exercise evaluation, made by the Exercise Control Group.

One of the main expected benefits of such a form of education is, in particular, improving the training of public administration representatives in the field of hybrid threats.
A positive aspect, however, is also raising awareness of hybrid threats in public administration, which will

ultimately lead to improved awareness of hybrid threats as such within the Slovak Republic, Visegrad Countries and the EU as a whole. For futher reading see also (Ristvej et al., 2016), (Gregor, et al. 2015), etc.

CONCLUSION

Nowadays, it is important to place increasing emphasis on the specific role of use of modelling and especially use of simulation in process of preparedness with focus on crisis management and resilience of the systems. Their importance and use is clearly confirmed at all phases of the cycle of crisis management.

On the other levels of crisis management, it is the lack of preparedness, whether of officials, operators or their superiors, as well as the unwillingness or rather ignorance of technologies based on the Open Source solutions – results of the project under the 7th FP, project COBACORE, etc.

Expected benefits are mainly connected to public administration officials training improving in the field of hybrid threats identifications, raising awareness of hybrid threats in public administration in Slovak Republic in general and improving awareness of hybrid threats within the Slovak Republic population.

Future development of this concept of using simulation for preparedness for hybrid threats is connected mainly to the recommendations complementing the training of public administration officials of hybrid threats. And expected resilience improvement of the public administration system to hybrid threats as a whole.

ACKNOWLEDGEMENT

"This publication was realized with support
of the Operational Program Research and Innovation
in frame of the project: ICT for smart society,
code ITMS2014 +: 313011T462, co-financed by
the European Regional Development Fund."

The views expressed, however, are solely those of the authors and not necessarily those of the institutions with which they are affiliated or of their funding sources. The authors are solely responsible for any errors or omissions.

REFERENCES

European Commission. 2016. "Joint Framework on countering hybrid threats: A European Union response." Available on: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52016JC0018&from=en

Government of the Slovak Republic. 2018. "Framework of the Slovak Republic on Countering Hybrid Threats." Available on: http://www.rokovania.sk/Rokovanie.aspx/BodRokovaniaDetail?idMaterial=27668

Gregor, M. et al. 2015. "Knowledge in healthcare " In: M.Annals of DAAAM Symposium. Volume 2015-January, 2015, Pages 1115-1121, Code 123540.

Loveček, T. 2008. "Present and future ways of Phisical Property Protection." In: Communications - Scientific Letters of the University of Žilina. Volume 10, Issue 1. 35-39. ISSN 1335-4205.

Meesters, K., Olthof, L. and B. Walle. 2014. "Disaster in my Backyard: a Serious Game to Improve Community Disaster Resilience." Researchgate. DOI: 10.13140/RG.2.1.3028.2084.

Ristvej, J. et al. 2016. "Simulation technologies in risk prevention within crisis management." In ESM´2016 = Modelling and simulation 2016: the European simulation and modelling conference 2016: October 26-28, 2016 Las Palmas, Gran Canaria, Spain. 327-330. ISBN 978-90-77381-95-3.

Rybár, M. et al. 2000. "Modelling and simulation in the military." [Modelovanie a simulácie vo vojenstve.] Bratislava: Publishing and Information Agency, Ministry of Defence of the Slovak Republic. 398 p. ISBN 80-88842-34-4.

Vichova, K., Hromada, M. and Rehak, D. 2017. "The Use of Crisis Management Information Systems in Rescue Operations of Fire Rescue Service of the Czech Republic. " Procedia Engineering, 12th International Scientific Conference Of Young Scientists On Sustainable, Modern and Safe Transport, TRANSCOM 2017; Volume 192, 2017, Pages 947-952

Walle, van de, B. et al. 2010. "Information Systems for Emergency Management." New York: ME. Sharpe, 2010, ISBN 978-0-7656-2134-4.

WEB REFERENCES

Raiszadeh B. and J. Batterson. 2018. "Live, Virtual, and Constructive Simulations in Test and Evaluation." In CK-12 Modeling and Simulation for High School Teachers: Principles, Problems, and Lesson Plans. Available at: https://www.ck12.org/book/CK-12-Modeling-and-Simulation-for-High-School-Teachers:-Principles-Problems-and-Lesson-Plans/section/4.2/

"Simultion – Type of simulation" [online]. 2019. Simulion Centre Armed Forces Academy Slovak Republic. Availabne on: http://www.aos.sk/sc/index.php?go=2.

Walker JA, Ruberg GE and JJ O'Dell. 1989. "Simulation for emergency management: taking advantage of automation in emergency preparedness." Simulation 53(3), 95–100.

MEDICAL SIMULATION

FINDING AN AGREEMENT WITH MIDPOINTS BETWEEN FUZZY SETS IN MEDICINE

Pilar Fuster-Parra
Universitat de les Illes Balears
Ctra. Valldemossa km. 7.5
07122-Palma de Mallorca, Spain
Institut d'Investigació Sanitária
Illes Balears (IdISBa)
Hospital Universitari Son Espases
07120 Palma de Mallorca
E-mail: pilar.fuster@uib.es

Javier Martín
Universitat de les Illes Balears
Ctra. Valldemossa km. 7.5
07122-Palma de Mallorca, Spain
Institut d'Investigació Sanitária
Illes Balears (IdISBa)
Hospital Universitari Son Espases
07120 Palma de Mallorca
E-mail: javier.martin@uib.es

Beatriz Romero-Ferrando
Waikato Hospital
Pembroke Street
3240 Hamilton
New Zeeland
E-mail: beatriz.romeroferrando@
waikatodhb.health.nz

Óscar Valero
Universitat de les Illes Balears
Ctra. Valldemossa km. 7.5
07122-Palma de Mallorca, Spain
Institut d'Investigació Sanitária
Illes Balears (IdISBa)
Hospital Universitari Son Espases
07120 Palma de Mallorca
E-mail: o.valero@uib.es

KEYWORDS

Hamming distance, decision making, midpoint, medical diagnosis.

ABSTRACT

In many real situations a practitioner must adopt an intermediate position between two fixed postures, that may even be opposite, in order to achieve an agreement and a working decision. In 2003, J.J. Nieto and A. Torres developed a mathematical technique for decision making in the aforementioned case by means of the use of the notion of midpoint between fuzzy sets. Such a technique establishes that the intermediate position can be adopted among the set of midpoints which is, in most cases, infinite (when the Hamming distance is considered) and, in addition, that the practitioner must decide from the infinite set of midpoints the right one. In order to help the practitioner, which is who has the art of finding the appropriate midpoint among all possible, to choose from the infinite set of midpoints in the decision making process, a decision making criterion based on the use of a scoring function on the set of midpoints is presented in such a way that the possible intermediate position between the two fixed ones is shortened to a finite number of midpoints at which the scoring function is either maximized or minimized.

As an illustrative example, we apply our decision making criterion to the case in which two independent and reputed doctors provide a patient description, one per doctor, and a medical committee must decide which description is more appropriate to be considered and from which to decide which is the final diagnosis.

INTRODUCTION

Fuzzy theory plays an important role in artificial intelligence, and has become a standard tool in some medical applications where uncertainty is present (Casasnovas and Rosselló 2004; Casasnovas and Rosselló 2005; Sadegh-Zadeh 1999; Sadegh-Zadeh 1999b). In Medicine, in many practical cases, when a symptoms profile of a patient is provided such symptoms are described by a finite number of numerical parameters whose values are fixed under certain uncertainty or ambiguity, because each numerical value represents the intensity or severity of a symptom shown by the patient and quantified by a doctor according to his/her medical experience. On account of (Nieto and Torres 2003), a suitable way to model such an information is using fuzzy sets in such a way that the symptoms profile can be described by a Kosko vector in the sense of (Kosko 1992).

When two reputed doctors provide independently two symptoms profile, represented as a Kosko vector, about the same patient a question arises in a natural way: what is the symptoms profile of both that must be considered for the medical diagnosis ? In (Nieto and Torres 2003), J.J. Nieto and A. Torres provided a suitable approach for this kind of decision making situations using the concept of midpoint between fuzzy sets. Concretely, they stablished that a practitioner which must make a decision about the diagnosis could take the consensus posture as a midpoint between both symptoms profiles that represent the doctors descriptions. However, the set of midpoints is, in general, infinite. In view of this situation, Nieto and Torres suggested that some additional criteria would be considered for the practitioner in order to decide a working midpoint description which allows

to make a final "middle way" diagnosis. So, a description of the midpoints between two fuzzy subsets has been formalized by Nieto and Torres when the Hamming distance is under consideration (Nieto and Torres 2003) and, in addition, given two fuzzy subsets, they give an infinite set of midpoints (intermediate positions or symptom profiles) to choose from and, finally, they leave to the practitioner the art of finding the right midpoint.

The aim of this work is to continue the original work of Nieto and Torres. Specifically we introduce a decision making criteria based on the use of a scoring function on the set of midpoints in such a way that the possible intermediate positions between two fixed ones is shortened to a finite number of midpoints at which the scoring function is either maximized or minimized. Besides, we apply our decision making criteria to the case in which two independent and reputed doctors provide a patient description, one per doctor, and a medical committee must decide which description is more appropriate to be considered and from which to decide which is the final diagnosis.

MATHEMATICAL BACKGROUND

To get deeper on this subject let us remember some basic concepts presented in (Nieto and Torres 2003). Fixed $n \in \mathbb{N}$ ($\mathbb{N}$ stands for the set of positive integer numbers) and a finite set $X = \{x_1, \ldots, x_n\}$, we will denote by $\mathcal{FP}(X)$ the set of its $[0,1]$-valued fuzzy subsets. Thus $\mathcal{FP}(X) = \{\mu : X \to [0,1]\}$. Note that every $\mu \in \mathcal{FP}(X)$ matches up with a vector $(\mu_1, \ldots, \mu_n) \in [0,1]^n$. So, according to (Kosko 1992), operations on the elements belonging to $\mathcal{FP}(X)$ can be performed on vectors in the Kosko hypercube $[0,1]^n$.

Following (Nieto and Torres 2003), the set of midpoints, $mid_d(\mu,\nu)$, between $\mu, \nu \in \mathcal{FP}(X)$, is defined as

$$mid_d(\mu,\nu) = \{\xi : d(\mu,\xi) = d(\xi,\nu) = \frac{1}{2}d(\mu,\nu)\},$$

where d is a distance defined on $\mathcal{FP}(X)$. Observe that the set $mid_d(\mu,\nu)$ depends on the distance d considered. Moreover, given $\mu, \nu \in \mathcal{FP}(X)$, the fuzzy subset ξ defined by $\xi(x_i) = \frac{1}{2}(\mu(x_i)+\nu(x_i))$ for all $i = 1, \ldots, n$ is a midpoint between μ and ν (called canonical midpoint in (Nieto and Torres 2003)). Furthermore, note that, in most cases, there is not a unique midpoint, i.e., the set $mid_d(\mu,\nu)$ is formed by infinite fuzzy sets. Clearly, this case is different from the case in which the distance d is exactly the Euclidean distance d_E, since there is only one midpoint, the canonical one, between two given fuzzy with respect to d_E.

Let us recall that, given two fuzzy subsets μ and ν, the fuzzy Hamming distance between μ and ν is defined by

$$d_H(\mu,\nu) = \Sigma_{i=1}^{n} |\mu(x_i) - \nu(x_i)|.$$

Nieto and Torres characterized the midpoint sets between fuzzy sets belonging to $\mathcal{FP}(X)$ when the 2-dimensional case and the Hamming distance d_H were considered. Concretely, the following result was proved in (Nieto and Torres 2003):

Theorem 1. *Let $X = \{x_1, x_2\}$ and let $\mu, \nu \in \mathcal{FP}(X)$. Then the following assertions hold:*

1. *If $\mu_1 = \nu_1$ or $\mu_2 = \nu_2$, then $\mathrm{mid}_{d_H}(\mu,\nu) = \{\frac{\mu+\nu}{2}\}$.*
2. *If $\mu_1 < \nu_1$ and $\mu_2 < \nu_2$, then $\mathrm{mid}_{d_H}(\mu,\nu) = \{\zeta \in \mathcal{FP}(X) : \mu_1 \le \zeta_1 \le \nu_1$ and $\zeta_2 = \frac{\mu_2+\nu_2+\mu_1+\nu_1}{2} - \zeta_1\}$.*
3. *If $\mu_1 < \nu_1$ and $\mu_2 > \nu_2$, then $\mathrm{mid}_{d_H}(\mu,\nu) = \{\zeta \in \mathcal{FP}(X) : \mu_1 \le \zeta_1 \le \nu_1$ and $\zeta_2 = \frac{\nu_2+\nu_1-\mu_1-\mu_2}{2} + \zeta_1\}$.*

It must be stressed that, in the light of Theorem 1 and in contrast to the Euclidean case, the set of midpoints between fuzzy sets does not reduce only to the canonical midpoint in general when the Hamming distance is employed. In fact, in most cases, the set of midpoints is an infinity set. This fact agrees with (Moderson et al. 2000), where it is stated that the working description (symptoms profile) can be identified with a wide range of "middle ways" between the original descriptions (symptoms profiles). A few applications of the Nieto and Torres approach to medical diagnosis can be found in (Nieto and Torres 2003). Notice, in addition, that regarding to the medical diagnosis problem aforementioned in Section , given the symptoms profiles, as Kosko vectors, that represent the original doctors descriptions of the patient, the consensus posture as a midpoint fails, in general, to be the canonical midpoint.

THE NEW DECISION MAKING CRITERIA

Motivated by the fact that set of midpoints is, in general infinite, and this causes that the practitioner must choose between infinite consensus postures, we introduce two decision making criteria which add to Nieto and Torres technique an additional criterion in order to reduce the number of consensus postures. Specifically the new decision making criterion is based on the use of a scoring function on the set of midpoints in such a way that the possible intermediate positions between the two original fixed ones is shortened to a finite number of midpoints at which the scoring function is either maximized or minimized.

In order to introduce the criterion, assume that we start from two postures which are described by two fuzzy sets $\mu, \nu \in \mathcal{FP}(X)$ defined on a finite set of attribute variables $X = \{x_1, \ldots, x_n\}$. Let us denote by $\mathbb{R}^+$ the set of non-negative real numbers. Consider, in addition, the continuous function $S : mid_{d_{H,\mathcal{FP}(X)}}(\mu,\nu) \to \mathbb{R}^+$ defined on the midpoints set $mid_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)$ of compromises or consensual positions between μ and ν by $S(\mu) = \sum_{i=1}^{n} \alpha_i \mu_i$ for all $\mu \in mid_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)$, where $\alpha_i \in \mathbb{R}_0^+$ for all $i = 1, \ldots, n$ ($\mathbb{R}_0^+$ stands for the positive real numbers set).

Then the numerical value $S(\mu)$ can be understood as the utility, scoring or payoff associated to the midpoint μ. Moreover, we are based on the assumption that the scoring value $S(\mu)$ is something in which larger values are worse and high values are not desirable. Thus the higher the value $S(\mu)$, the lower the utility associated to μ. Moreover, each α_i can be understood as a way to express the importance of the value μ_i in

order to make the working decision. Clearly we avoid the value 0 for any α_i, since otherwise we merge a value of an attribute variable which does not matter to make the decision. Furthermore, it is not hard to check that S attains the (global) minimum and (global) maximum values at an extreme point of $mid_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)$ and that the aforesaid number of extreme points is finite, since the scoring function S is linear and continuous and the midpoints set $mid_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)$ is a compact polyhedron (see, (Luc 2016) and (Ok 2007)).

Let us denote by $Min_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)$ and $Max_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)$ the midpoint sets such that S attains its minimum and maximum value, respectively. Then, according to the previous discussion $Min_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)\neq\emptyset$, $Max_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)\neq\emptyset$ and, in addition, we have guaranteed the existence of two points $\zeta_{min}\in Min_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)$ and $\zeta_{max}\in Max_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)$ such that S attains its (global) minimum value at ζ_{min} and its (global) maximum value at ζ_{max}.

Then the new decision making criterion proposes, in order to reduce the midpoint set to a finite number of working midpoints, The Minimizer/Maximizer Criterion as decision rule.

The Minimizer Criterion (Maximizer Criterion) can be understood as an optimistic (pessimistic) approach in the sense of (French 1988). Notice that the decision maker (practitioner) ranks all midpoints in based on their scoring values and he/she selects the midpoint at which the scoring function attains its minimum (maximum) value giving the greatest utility (lowest utility), since high numerical values $S(\mu)$ are not desirable. Observe that although the proposed criteria yield an assisted decision rule, it also gives a degree of freedom to the decision maker, since it allows the decision maker to choose, based on the acquired experience, the most appropriate midpoint among all midpoints belonging to either $Min_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)$ or $Max_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)$. A very interesting case is given when the cardinality either $Min_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)$ or $Max_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)$ is one, that is, either $Min_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)=\{\zeta_{min}\}$ or $Max_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)=\{\zeta_{max}\}$. In this last case, the minimizer criterion (maximizer criterion) provides an assisted decision rule which is completely directed until the decision making. This is of special interest in those cases in which the decision maker needs that the decision rule provides a final decision in such a way that the his/her unique intervention is either to validate the final working decision or to take it into account as a starting point to make a decision, always based on his/her experience acquired.

CHARACTERIZING EXTREME MIDPOINTS

In this section we provide a description of those midpoints that are also extreme points of the scoring function when the cardinality of the set of attributes is exactly two. To this end, let us consider $X=\{x_1,x_2\}$ and $\alpha_1,\alpha_2\in\mathbb{R}^{++}$. Then the scoring function is given by $S(\mu)=\alpha_1\mu_1+\alpha_2\mu_2$. Assume that $\mu,\nu\in\mathcal{FP}(X)$. It is enough to consider only the following two possible cases:

Case i). $\mu_1<\nu_1$ and $\mu_2<\nu_2$,

Case ii). $\mu_1<\nu_1$ and $\mu_2>\nu_2$.

We avoid the case in that $\mu_1=\nu_1$ or $\mu_2=\nu_2$ because then the consensus between μ and ν is achieved at the canonical midpoint $\frac{\mu+\nu}{2}$ and the unique extreme point is exactly the canonical one.

Next we assume that the Case i) is hold. Theorem 1 gives the following midpoint set

$$\mathrm{mid}_{d_H}(\mu,\nu)=\{\zeta\in\mathcal{FP}(X):\mu_1\leq\zeta_1\leq\nu_1 \text{ and } \zeta_2=\frac{\mu_2+\nu_2+\mu_1+\nu_1}{2}-\zeta_1\}.$$

Let us assume that $\nu_1-\mu_1\leq\nu_2-\mu_2$. Then it is not hard to check that the set of extreme points of the polyhedron $\mathrm{mid}_{d_H}(\mu,\nu)$ is exactly $\{(\mu_1,\frac{\nu_2+\mu_2+\mu_1-\nu_1}{2}),(\nu_1,\frac{\nu_2+\mu_2+\nu_1-\mu_1}{2})\}$.

The next result obtains the minimizer and maximizer midpoints of the scoring function.

Proposition 2. *Let $X=\{x_1,x_2\}$, $\mu,\nu\in\mathcal{FP}(X)$ such that $\nu_1-\mu_1\leq\nu_2-\mu_2$. Consider the function $S:\mathrm{mid}_{d_H}(\mu,\nu)\to\mathbb{R}^+$ given by $S(\mu)=\alpha_1\mu_1+\alpha_2\mu_2$, where $\alpha_1,\alpha_2\in\mathbb{R}^+_0$. Then the following assertions hold:*

1. *If $\alpha_2>\alpha_1$, then we have that $Max_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)=\{(\mu_1,\frac{\nu_2+\mu_2+\nu_1-\mu_1}{2})\}$ and $Min_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)=\{(\nu_1,\frac{\nu_2+\mu_2+\mu_1-\nu_1}{2})\}$.*
2. *If $\alpha_1>\alpha_2$, then we have that $Max_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)=\{(\nu_1,\frac{\nu_2+\mu_2+\mu_1-\nu_1}{2})\}$ and $Min_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)=\{(\mu_1,\frac{\nu_2+\mu_2+\nu_1-\mu_1}{2})\}$.*
3. *If $\alpha_2=\alpha_1$, then we have that $Max_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)=Min_{d_{H,\mathcal{FP}(X)}}(\mu,\nu)=\{(\mu_1,\frac{\nu_2+\mu_2+\nu_1-\mu_1}{2}),(\nu_1,\frac{\nu_2+\mu_2+\mu_1-\nu_1}{2})\}$.*

Notice that in case that $\nu_1-\mu_1>\nu_2-\mu_2$ all computations suggested by our criteria can be made following Proposition 2 when we interchange the fuzzy sets μ and ν by μ^* and ν^* such that $\mu_1^*=\mu_2$, $\mu_2^*=\mu_1$, $\nu_1^*=\nu_2$ and $\nu_2^*=\nu_1$. Of course the preceding suggested change implies that we have changed the order of the attributes variables in the set X and, thus, the above possible conclusions provided by our criteria remain valid.

Next we assume that the Case ii) is hold. Theorem 1 gives the following midpoint set

$$\mathrm{mid}_{d_H}(\mu,\nu)=\{\zeta\in\mathcal{FP}(X):\mu_1\leq\zeta_1\leq\nu_1 \text{ and } \zeta_2=\frac{\mu_2+\nu_2-\mu_1-\nu_1}{2}+\zeta_1\}.$$

Let us assume that $\nu_1-\mu_1\leq\mu_2-\nu_2$. Then it is not hard to check that the set of extreme points of the polyhedron $\mathrm{mid}_{d_H}(\mu,\nu)$ is exactly $\{(\mu_1,\frac{\nu_2+\mu_2+\mu_1-\nu_1}{2}),(\nu_1,\frac{\nu_2+\mu_2+\nu_1-\mu_1}{2})\}$.

The next result obtains the minimizer and maximizer midpoints of the scoring function.

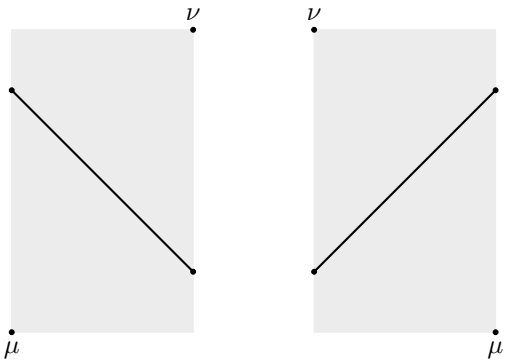


Figure 1: Midpoints and extremes obtained from the fuzzy subsets $\mu = (\mu_1, \mu_2)$ and $\nu = (\nu_1, \nu_2)$, on the left $\nu_1 - \mu_1 \leq \nu_2 - \mu_2$, on the right $\nu_1 - \mu_1 < \mu_2 - \nu_2$.

Proposition 3. *Let $X = \{x_1, x_2\}$, $\mu, \nu \in \mathcal{FP}(X)$ such that $\nu_1 - \mu_1 \leq \mu_2 - \nu_2$. Consider the function $S : \mathrm{mid}_{d_H}(\mu, \nu) \to \mathbb{R}^+$ given by $S(\mu) = \alpha_1\mu_1 + \alpha_2\mu_2$, where $\alpha_1, \alpha_2 \in \mathbb{R}_0^+$. Then $Max_{d_{H,\mathcal{FP}(X)}}(\mu, \nu) = \{(\nu_1, \frac{\nu_2+\mu_2+\nu_1-\mu_1}{2})\}$ and $Min_{d_{H,\mathcal{FP}(X)}}(\mu, \nu) = \{(\mu_1, \frac{\nu_2+\mu_2+\mu_1-\nu_1}{2})\}$.*

Observe, again, that in case that $\nu_1 - \mu_1 > \mu_2 - \nu_2$ all computations suggested by our criteria can be made following Proposition 3 when we interchange the fuzzy sets μ and ν by μ^* and ν^* such that $\mu_1^* = \mu_2$, $\mu_2^* = \mu_1$, $\nu_1^* = \nu_2$ and $\nu_2^* = \nu_1$.

Figure 1 illustrates the set of midpoint and the set of extremes when $\nu_1 - \mu_1 \leq \nu_2 - \mu_2$, and $\nu_1 - \mu_1 < \mu_2 - \nu_2$.

It must be pointed out that the minimizer and maximizer midpoints provided by Proposition 3 are independent of the values taken by the coefficients α_1 and α_2. This is a remarkable difference with respect to the Case $i)$ which has been exposed in Proposition 2.

AN ILLUSTRATIVE EXAMPLE FROM MEDICINE: ALZHEIMER VERSUS VASCULAR DEMENTIA

In this section we present an example in order to clarify our decision making criteria and to show how it can be helpful in medicine. In the particular case of medical diagnosis, the set $X = \{x_1, x_2\}$ can be interpreted as two illnesses (Alzheimer and Vascular Dementia) which have common symptoms and they are not easy to differentiate at an early stage by a doctor as many patients have vascular risk factors and small vessel disease can be found in the brain imaging. It is known that a quick diagnosis would help the patients and their families to understand the disease, be aware of possible complications and know the prognosis. Notice that when a doctor evaluate the symptoms profile of the patient and try to quantify the intensity of the illnesses (Alzheimer versus Vascular Dementia) is shown by the patient, there exists a certain uncertainty or ambiguity, because such a quantification is made by a doctor according to his/her medical experience. So it is very usual that two doctors evaluate the same patient (which

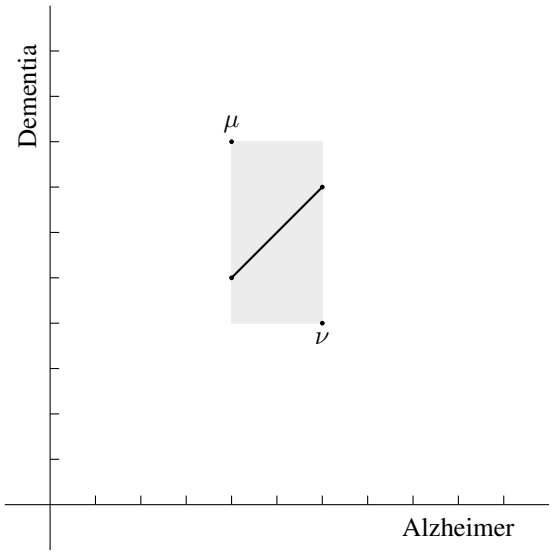


Figure 2: Midpoints and extremes obtained from the fuzzy subsets $\mu = (0.4, 0.8)$ and $\nu = (0.6, 0.4)$.

is a typical situation when a patient asks for a second medical opinion) and they differ in their opinions. So a medical committee must decide which is the final diagnosis. Our criteria will allow to make a consensus via a scoring function on the midpoint sets.

Let us assume that Doctor 1 thinks that the patient has Alzheimer in a 40% of possibilities and Vascular Dementia in 80% of possibilities. Doctor 2 thinks that the patient has Alzheimer in a 60% of possibilities and Vascular Dementia in 40% of possibilities.

Both practitioners have enough medical experience and reputation, so both diagnosis must be taken into consideration. Observe that the fuzzy sets are $\mu = (0.4, 0.8)$ and $\nu = (0.6, 0.4)$. Notice, in addition, that a description of the type $(1, 0)$ and $(0, 1)$ represents that the doctor is convinced that the patient suffers Alzheimer and Vascular Dementia, respectively.

Next let us consider a scoring function $S : mid_{d_{H,\mathcal{FP}(X)}}(\mu, \nu) \to \mathbb{R}^+$ given by $S(\mu) = 0.9\mu_1 + 0.6\mu_2$. Notice that the weights are suggesting that Alzheimer have more influence than vascular Dementia in order to make the final diagnosis by a committee. According to Proposition 3 $Max_{d_{H,\mathcal{FP}(X)}}(\mu, \nu) = \{(\nu_1, \frac{\nu_2+\mu_2+\nu_1-\mu_1}{2})\} = \{(0.6, 0.7)\}$ and $Min_{d_{H,\mathcal{FP}(X)}}(\mu, \nu) = \{(\mu_1, \frac{\nu_2+\mu_2+\mu_1-\nu_1}{2})\} = \{(0.4, 0.5)\}$.

According to the above situation, consensus postures between both original evaluations from the independent doctors can be obtained by the means of the midpoint set $mid_{d_{H,\mathcal{FP}(X)}}(\mu, \nu)$. Notice that there are infinity midpoints, since $mid_{d_{H,\mathcal{FP}(X)}}(\mu, \nu) = \{(\zeta_1, \zeta_2) \in [0, 1]^2 : 0.4 \leq \zeta_1 \leq 0.6 \text{ and } \zeta_2 = 0.1 + \zeta_1\}$. So, in oder to reduce such a number of possible consensus postures, we can apply our new criteria obtaining the following one:

1. $(0.4, 0.5)$ according to the minimizer criterion (optimistic criterion),

2. $(0.6, 0.7)$ according to the maximizer criterion (pessimistic criterion).

Observe that the preceding consensus postures merge the viewpoint of both doctors trying to incorporate to the middle way description the quantitative information coming from both reputed doctors. Next a medical committee must decide, based on the experience acquired, the final diagnosis taking as a starting point one description, either the given by the minimizer criterion or that given by the maximizer criterion.

CONCLUSIONS

A decision making criterion based on the use of midpoints between fuzzy sets (in the sense of Nieto and Torres) and a scoring function has been introduced in such a way that the consensus posture between two fixed ones is reduced to a finite number of midpoints at which the scoring function is either maximized (Maximizer Criterion) or minimized (Minimizer Criterion). A numerical example in which a description of those midpoints that are also extreme points of the scoring function has been exposed. Finally, an example has been considered where two independent and reputed doctors provide two patient diagnoses, one per doctor, and the medical staff (another doctor or a committee of doctors) must decide which diagnosis is more appropriate to be considered and from which to decide which is the final diagnosis.

ACKNOWLEDGEMENTS

P. Fuster-Parra, J. Martín and O. Valero acknowledge the financial support from Spanish Ministry of Science, Innovation and Universities under Grant PGC2018-095709-B-C21 and AEI/FEDER, UE funds. This work is also partially supported by Programa Operatiu FEDER 2014-2020 de les Illes Balears, by project PROCOE/4/2017 (Direcció General d'Innovació i Recerca, Govern de les Illes Balears) and by project ROBINS. The latter has received research funding from the EU H2020 framework under GA 779776. This publication reflects only the authors views and the European Union is not liable for any use that may be made of the information contained therein.

REFERENCES

J. Casasnovas, F. Roselló, Midpoints as average representations of pairs of descriptions by means of fuzzy subsets, In Proc. Information Processing and Management of Uncertainty in Knowledge-based Systems international Conference (Edited by B. Bouchon-Meunier et al.), pp. 2157-2164. Editrice Universita La Sapienza. Roma, (2004).

J. Casasnovas, F. Rosselló, Averaging fuzzy biopolymers, Fuzzy Sets and Systems 152 (2005), 139-158.

S. French, Decision Theory: An Introduction to the Mathematics of Rationality, Ellis Horwood, 1988.

D.T. Luc, Multiobjective Linear Programming: An Introduction, Springer, 2016.

B. Kosko, Neural networks and fuzzy systems, Prentice-Hall, Englewood Clifss, 1992.

E.A. Ok, Real Analysis with Economic Applications, Princeton University Press, 2007.

J. Moderson, D. Malik, S. C. Cheng, Fuzzy Mathematics in Medicine, Physica-Verlag, 2000.

J.J. Nieto, A. Torres, Midpoints for fuzzy sets and their application in medicine, Artificial Intelligence in Medicine 27 (2003), 81-101.

K. Sadegh-Zadeh, Advances in fuzzy theory, Artificial Intelligence in Medicine 15 (1999), 309-323.

K. Sadegh-Zadeh, Fundamentals of clinical methodology:3. Nosology, Artificial Intelligence in Medicine 17 (1999), 87-108.

BIOGRAPHIES

PILAR FUSTER-PARRA is Associate Professor at the Universitat de les Illes Balears (Departament de Ciències Matemàtiques i Informàtica). She received her M.Sc. in Mathematics from Universitat de València in1988 and a Ph.D. in Computer Science from Universitat de les Illes Balears in 1996. Her research interests lie primarily in the areas of mathematical modelling, Artificial Intelligence, fuzzy sets, possibility theory, generalized distances, multi-agent systems.

JAVIER MARTIN is Professor at the Universitat de les Illes Balears (Departament de Ciències Matemàtiques i Informàtica) in the area of Computing Sciences and Artificial Intelligence. He received his M.Sc. in Mathematics from Universidad de Valladolid in1997. His research activity focus on generalized distances, aggregation operators and Fuzzy logic.

BEATRIZ ROMERO Beatriz Romero-Ferrando is a neurologist currently working at the Waikato Hospital. She received her M.Sc in Medicine and Surgery from Universitat de València in 2003. She obtained her Neurology Diploma from Hospital Nuestra Señora de Valme in 2008. She worked as a general neurologist in Hospital Son Llatzer, while she was focusing in Multiple Sclerosis and Epilepsy from 2008 to 2019.

OSCAR VALERO is Associate Professor at Universitat de les Illes Balears (Departament de Ciències Matemàtiques i Informàtica). He received his M.Sc. in Mathematics from

Universitat de València in 2000. Later on, in 2003, he obtained a Ph.D. in Mathematics from Universitat Politècnica de València in the Doctoral Program of Multidisciplinary Mathematics. In addition, in 2003 he made a short stay of research in the "Center for Efficiency Oriented Languages" of the Department of Computer Science of the National University of Ireland. He is currently part of the center as an "Affiliate Member" researcher. His research activity focus on Mathematical Models applied to Computer Science, Artificial Intelligence, Engineering, Economics and Medicine.

Using Machine Learning and Heart Rate Variability Features to Predict Epileptic Seizures

Antoni Burguera
Departament de Matemàtiques i Informàtica
Universitat de les Illes Balears
Carretera de Valldemossa Km 7.5 - 07122 Palma
Spain
email: antoni.burguera@uib.es

KEYWORDS
Machine learning, biomedical informatics, electrocardiogram, epilepsy prediction

ABSTRACT

This study constitutes a first step towards a wearable epileptic seizure prediction device. We exploit the existing correlation between epileptic pre-ictal states and heart rate variability features, since they can be measured by portable electrocardiogram recorders. By explicitly dealing with the intervals of extreme noise that may corrupt the electrocardiogram data during the seizures, our proposal is able to robustly train and use a Support Vector Machine to detect pre-ictal states. The experimental results show particularly good results in terms of positive and negative prediction. They also show the importance of a specific training for each patient.

Introduction

Although epileptic seizures are controllable with medication in 70% of the cases (Eadie 2012), the remaining 30% can lead to collateral physical injuries. Taking into account that approximately 1% of the world population is affected by epilepsy, creating a wearable technology able to predict epileptic seizures and warn the patient before they happen would have an immense social impact.

Most of the existing studies focused on epileptic seizure prediction (Gadhoumi, K. et al 2016) rely on the data provided by an electroencephalogram (EEG). Consequently, they cannot be used in daily life.

Some studies (Al-Aweel et al. 1999, Fujiwara, K. et al. 2015) have found statistical correlation between heart rate variability (HRV) features and epileptic pre-ictal and post-ictal states. Even though the underlying physiological mechanisms are not fully understood, the fact that these correlations exist opens a door to epileptic seizure prediction by means of electrocardiography (ECG).

The main advantage of ECG in front of EEG is its portability (Baig, M.M. et al. 2013), making it possible to build wearable predictors. This is, precisely, the goal of this paper: to define a methodology to reliably predict epileptic seizures from the data provided by portable ECG recorders. This goal is also one of the main novelties of this study, since few researchers (Hashimoto, H. et al. 2013) have addressed the problem of epileptic seizure prediction from ECG.

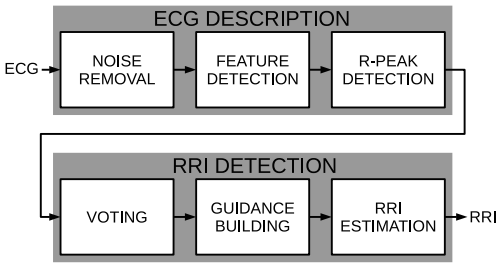


Figure 1: ECG processing overview.

More specifically, this paper focuses on the use of supervised classification methods to detect pre-ictal HRV patterns and predict epileptic seizures. In particular, the paper begins by presenting a fast and robust approach to process ECG data. This approach not only deals with the noisy nature of portable ECG recorders but also with the intervals of severe noise that may appear during the seizures. Afterwards, we discuss the HRV features used to perform the prediction. These features are subsequently used to train the Support Vector Machine (SVM) in charge of achieving seizure prediction, which is also described. Finally, the paper presents several experiments conducted using real ECG data gathered before, during and after epileptic seizure episodes.

ECG processing

The goal of the ECG processor is to robustly estimate the *inter-beat intervals*, which are defined as the time between consecutive heart beats. A common approach to obtain them is to measure the time between consecutive *R-peaks* in an ECG. Because of that, the inter-beat intervals are also known as the *R-R intervals* or RRI.

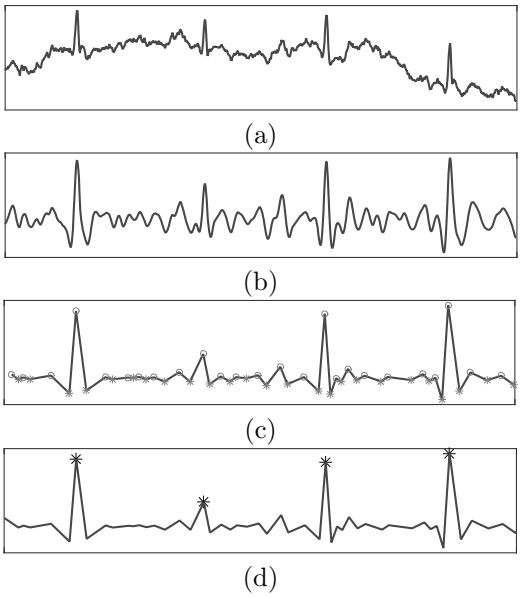


Figure 2: (a) Example of raw ECG (b) Noise removal (c) Feature detection (d) R-peak detection

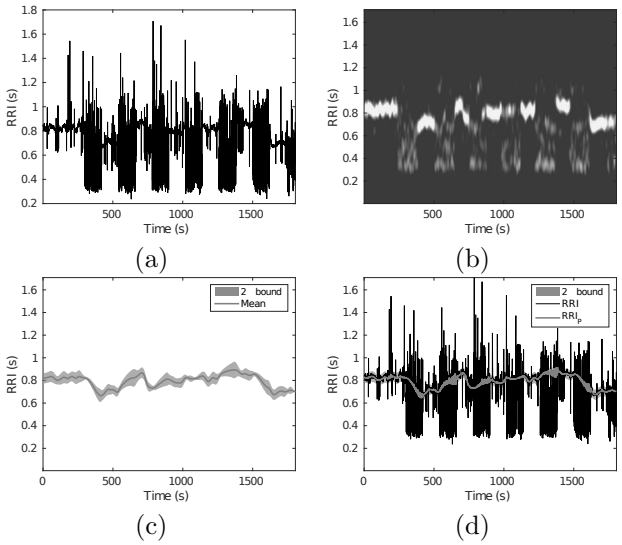


Figure 3: (a) Initial RRI (b) Voting table (c) Guidance signal (d) Final RRI

Even though several approaches addressing this problem exist in the literature (Elgendi, M. et al. 2014), our proposal has some important advantages: it is fast and parameterless (Burguera 2018) and it explicitly deals with the intervals of extreme noise that may appear during epileptic seizures.

Our proposal performs two main steps, depicted in Figure 1. The first step, *ECG description*, processes the raw ECG and detects the existing R-peaks. The second step, *RRI detection*, robustly estimates the RRI.

ECG description

An ECG is influenced by a wide range of noise sources. For example, breathing or muscular activity are responsible for low frequency noise and cause baseline drift whilst external electrical interferences produce high frequency noise. Figure 2-a shows an ECG influenced by these noise sources. In order to deal with them, the *noise removal* step is first performed.

Even though band-pass filters are a common choice to deal with this problem some studies (Burguera 2018) propose high-pass filtering as a better option for two main reasons. On the one hand, some artifacts and ripples may appear due to the low-pass components. On the other hand, it may be difficult to select a cutoff frequency that does not interfere with relevant ECG waves. Accordingly, our proposal is to detrend the ECG (i.e. remove the baseline drift) by means of a high-pass filter and, afterwards, remove the high frequency noise by means of a signal smoothing method. As for detrending, the high-pass filter is implemented as a Butterworth *Infinite Impulse Response* (IIR) filter of order three with zero-phase distortion and a cut-off frequency of 5 Hz.

Smoothing is achieved by means of the so called *Moving Linear Regression* (MLR), which averages the ECG within a sliding window using of a set of best fitting lines. A detailed description as well as a closed-form formulation can be found in (Burguera 2018). Figure 2-b shows the output of the noise removal step.

Once the noise has been removed, our proposal is to search the peaks and the valleys within the denoised data as they preserve the main ECG structure. The set of peaks and valleys is then refined by removing those which are too close to belong to actual heart activity. This constitutes the *feature detection* step, which provides the set of key points to be subsequently used to detect the R-peaks. Figure 2-c exemplifies this step.

R-peak detection is achieved by taking into account the whole QRS complex structure. A QRS complex is a well known succession of waves within an ECG that usually follows a particular and easily identifiable structural pattern (Pandit, D. et al. 2017). The subsets of consecutive features that meet this pattern are searched to build an initial estimate of the existing QRS complexes. Additionally, in order to deal with local ECG amplitude changes, the signal envelope is continuously computed (Burguera 2018) and used to improve the pattern matching results. Afterwards, the initial estimate is iteratively refined by removing those complexes which are too close to be feasible and by accurately searching missing complexes between those which are too separated.

The R-peaks within the detected QRS complexes, illustrated in Figure 2-d, constitute the output of the ECG description step, and will be used to robustly estimate the RRI. A ready to use implementation of the ECG descriptor is available at `https://github.com/aburguera/QRS_DETECT`

RRI detection

The RRI can be computed by subtracting the time steps of consecutive heart beats. However, intervals of extreme noise may appear due to the epileptic seizures. During these intervals, it is difficult to detect the R-peaks even for human experts. This section proposes a novel approach to improve the RRI estimates during these intervals of extreme noise.

The RRI detection operates in the three steps shown in Figure 1. The *voting* step begins by building an initial RRI estimate. This estimate is built by subtracting the time steps of consecutive R-peaks. Figure 3-a shows an example of initial RRI corresponding to an ECG with intervals of extreme noise. After building the initial RRI, a voting table is created and each initial RRI estimate votes those RRI values that are acceptable from its own point of view. As a result, the votes concentrate around the correct values whilst wrong RRI, which tend to be isolated, receive less votes. Figure 3-b exemplifies the result of this step.

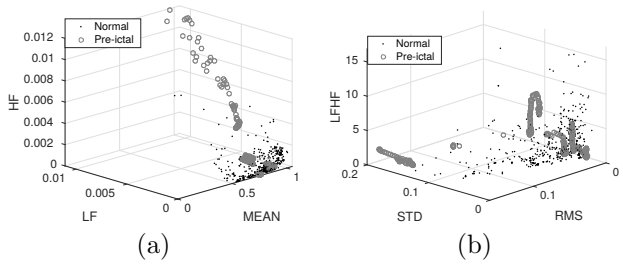
The *guidance building* step applies a dynamic threshold to the voting table in order to distinguish the clear and the noisy regions. Then, a guidance signal is built so that it follows the clear regions and is interpolated in the noisy regions. The voting table is also used to determine a confidence interval for this guidance signal. Figure 3-c shows an example of a guidance signal constructed in this way.

Finally, the *RRI estimation* step fuses the initial RRI and the guidance signal by means of a Kalman filter (KF). As a result an improved RRI (RRI_P) is constructed, as illustrated in Figure 3-d. A ready to use implementation of the RRI detector is available at `https://github.com/aburguera/RRI_PROCESSOR`

HRV features

Different studies have observed a correlation between HRV features and pre-ictal states. In particular, (Fujiwara, K. et al. 2015) detected a significant correlation with set of five time domain features and three frequency domain features. The time domain features are the mean (MEAN), the standard deviation (STD), the root mean square of differences between adjacent values (RMS) and the total power (TPW). The frequency domain features are the power of the power spectrum density (PSD) band 0.04Hz-0.15Hz (LF), the power of the PSD band 0.15Hz-0.4Hz (HF) and the ratio between these features (LFHF). Our proposal is to use these features to describe the RRI. All these features are computed within a sliding window over the RRI estimates whose size will be experimentally assessed.

In Figure 4, some of the aforementioned HRV features have been plotted against each other and labelled as pre-ictal and normal by means of the ground truth. As it can be observed, even though three features are not sufficient to distinguish between pre-ictal and normal states, some correlation can be observed.



(a) (b)

Figure 4: Examples of HRV features and their correlation with pre-ictal states. (a) Mean RRI (MEAN), power of the PSD band 0.04Hz-0.15Hz (LF) and power of the PSD band 0.15Hz-0.4Hz (HF). (b) Root mean square of differences between adjacent RRI (RMS), standard deviation of the RRI (STD) and ratio between the powers of low and the high bands of the PSD (LFHF).

Supervised classification

Our proposal is to use the aforementioned HRV features to train a SVM with a radial basis function (RBF) kernel. Once trained, new data can be classified as pre-ictal or normal. In other words, epileptic seizures could be predicted on-line using the ECG data.

In order to train the system, a ground truth is necessary. In most cases, the ground truth will identify the seizure intervals instead of the pre-ictal ones. This raises the need for a new parameter to be tuned: the time window before a seizure in which we want to do the prediction. Let this parameter be referred to as the *prediction interval*.

Selecting a prediction interval is not trivial. On the one hand, a small interval may render the prediction useless, even if it is correct. Moreover, small prediction intervals also reduce the number of usable samples to compute the HRV features, jeopardizing the SVM performance. On the other hand, large intervals may lead to a wrong SVM training, since features within the large interval will include both normal and true pre-ictal situations.

Accordingly, the length of prediction interval as well as the size of the sliding window used to compute the HRV features will be experimentally assessed in the next section.

Experimental results

In order to evaluate our proposal, the SZDB database (Al-Aweel et al. 1999, Goldberger, A.L. et al. 2000) has been used. This database, composed of 7 records, contains more then 16 h of ECG sampled at 200 Hz. The data corresponds to five women patients, ranged in age from 31 to 48 years old, during continuous EEG, ECG and video monitoring. The recording involves 11 epilep-

Table 1: Monte-Carlo cross validation results. The table shows the mean (μ) and the standard deviation (σ) of the true and false positives and negatives for prediction intervals (PI) ranging from 1 to 5 minutes and for the two labelling criteria (binary and majority). In each case, the first pair μ/σ corresponds to the classification results if the RRI detection is not used whilst the second pair shows the results if it is used.

	Binary				Majority			
PI	TP (μ/σ)	TN (μ/σ)	FP (μ/σ)	FN (μ/σ)	TP (μ/σ)	TN (μ/σ)	FP (μ/σ)	FN (μ/σ)
1	4.2/3.1	51276.8/1.1	0.4/1.1	197.6/3.1	41.2/8.2	50944.5/3.2	2.5/3.2	490.8/8.2
	1.0/1.3	51277.0/0.6	0.2/0.6	200.8/1.3	8.0/3.8	50945.7/2.5	1.3/2.5	524.0/3.8
2	42.0/10.3	50795.5/4.7	3.1/4.7	386.4/10.3	228.0/20.7	50160.4/8.3	15.6/8.3	823.0/20.7
	50.9/11.8	50797.0/1.9	1.6/1.9	377.4/11.8	179.6/30.2	50165.1/8.3	10.9/8.3	871.4/30.2
3	123.9/26.9	50299.2/8.7	4.7/8.7	548.2/26.9	599.1/44.7	49380.5/25.4	28.5/25.4	967.9/44.7
	188.7/15.7	50303.2/1.4	0.7/1.4	483.4/15.7	504.0/39.5	49385.1/17.7	23.9/17.7	1063.0/39.5
4	423.4/9.8	49795.8/4.0	2.5/4.0	517.2/9.8	1099.6/19.2	48631.3/8.5	11.7/8.5	996.4/19.2
	478.5/23.4	49787.6/5.5	10.7/5.5	462.2/23.4	1172.0/37.9	48598.6/15.0	44.4/15.0	924.0/37.9
5	608.4/16.7	49292.0/4.3	9.3/4.3	607.4/16.7	1434.0/45.1	47871.6/11.9	28.4/11.9	1183.0/45.1
	650.9/22.2	49296.5/7.8	4.7/7.8	564.9/22.2	1383.0/35.2	47832.6/16.3	67.4/16.3	1234.0/35.2

Table 2: Results of training with each record R0 and classifying each record R1.

		PI=1 min							PI=3 min							PI=5 min						
	R0 \ R1	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
HR	1	99	99	99	99	99	98	99	100	97	97	97	96	93	97	98	95	95	94	94	93	96
	2	99	99	99	99	99	99	99	97	98	97	97	96	96	97	95	97	95	94	94	94	96
	3	99	99	99	98	99	99	99	97	97	98	91	96	96	97	95	94	98	94	94	94	96
	4	99	99	99	99	99	99	99	97	96	97	98	96	96	97	95	92	95	98	94	94	96
	5	99	99	99	99	99	99	99	97	97	97	97	99	96	97	95	95	95	94	97	94	96
	6	99	99	99	99	99	99	99	97	97	97	97	96	98	97	95	95	95	94	94	96	96
	7	99	99	99	99	99	99	99	96	96	97	97	96	96	98	95	95	95	94	94	93	99
SE	1	46	0	0	0	0	0	0	100	0	0	0	0	0	0	60	0	0	0	0	0	0
	2	0	25	0	0	0	0	0	0	46	0	0	0	0	0	0	48	0	0	0	0	0
	3	0	0	9	0	0	0	0	0	0	25	0	0	0	0	0	0	54	0	0	0	0
	4	0	0	0	27	0	0	0	0	0	0	51	0	0	0	0	0	0	63	0	0	0
	5	0	0	0	0	0	0	0	0	0	0	0	70	0	0	0	0	0	0	50	0	0
	6	0	0	0	0	0	0	0	0	0	0	0	0	30	0	0	0	0	0	0	26	0
	7	0	2	0	0	0	0	27	0	0	0	0	0	0	13	0	0	0	0	0	0	72
SP	1	100	100	100	100	100	99	100	100	100	100	100	100	97	100	100	100	100	100	100	99	100
	2	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	3	100	100	100	99	100	100	100	100	100	100	95	100	100	100	100	100	100	100	100	100	100
	4	100	100	100	100	100	100	100	100	99	100	100	100	100	100	100	97	100	100	100	100	100
	5	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	6	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	7	100	100	100	100	100	100	100	99	100	100	100	100	99	100	100	100	100	100	100	99	100

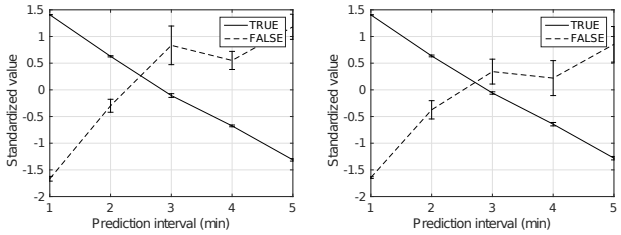
tic seizures lasting from 15 to 110 seconds. The EEG and video records were used by expert neurologists to tag the epileptic seizures. Thus, the ground truth is available and perfectly defined within this database.

In all the experiments, 5 prediction intervals have been tested, ranging from 1 minute to 5 minutes. The size of the sliding window used to compute the HRV features was set to half the prediction interval. Also, two methods to provide a ground truth for a given feature have been used. The first method, named *binary*, tags a feature as pre-ictal if all the samples within the corresponding sliding window are pre-ictal. The second method, named *majority* tags a feature as pre-ictal if the majority of samples within the corresponding sliding window are pre-ictal.

Afterwards, two experiments have been performed for each of these 10 configurations. In all cases, all the ECG have been processed as described, and the RRI as well as the HRV features have been obtained. The features corresponding to the seizure intervals themselves as well a few minutes following the seizure have been removed, so that the remaining data only involves pre-ictal and normal states.

The first experiment is a Monte-Carlo cross validation involving all the records. In particular, a 25% of the samples is randomly selected without replacement and used to train the SVM. The remaining 75% constitutes the test set and has been classified by the trained SVM. The quality of the classified samples has been assessed by means of true and false positives and negatives. This cross-validation schema, involving random training and test sets has been repeated 100 times for each configuration.

Table 1, which summarizes the results, shows that increasing the prediction interval leads to an increase in the number of true positives but also to less true negatives, resulting in an overall increment of false predictions.



(a) (b)

Figure 5: Standardized number of true and false detections as a function of the prediction interval if RRI detection is (a) not used and (b) used.

When comparing the labelling criteria, it can be observed that even though the majority criterion leads to an increase in the number of true positives, the negative effects in the remaining values are more relevant. Thus, the binary criterion is the best choice to train the SVM. Finally, the benefits of using the RRI detector can also be observed.

Figure 5 illustrates these effects. In particular, it can be observed how the number of true estimates decreases with the prediction interval while the number of errors increases. Additionally, by comparing Figures 5-a and 5-b it can be observed how the RRI detector barely influences the number of true estimates but it significantly reduces the number of wrong predictions, making it possible to predict seizures with more anticipation.

The second experiment involves training the SVM with each of the 7 records and then classifying and evaluating all the records, including the one used for training. The results for three prediction intervals (1, 3 and 5 min) are shown in Table 2 in terms of hit ratio (HR), sensitivity (SE) and specificity (SP). As it can be observed, the hit ratio and the specificity are very close to 100% in all cases but the sensitivity results are not so good. This difference is partially due to the large amount of negatives in all the records.

Also, sensitivities are very close to zero when the tested record is not the one used to train the system. This suggests a strong dependence on the specific patient. That is, specific training for each patient is needed.

Conclusion

This study constitutes a first step towards a wearable epileptic seizure prediction device. To this end, the input data is obtained from ECG, contrarily to other strategies which depend on EEG. Our proposal begins by processing the ECG explicitly dealing with the intervals of extreme noise that may appear during the epileptic seizures. Then, the processed data is used to train the SVM that will subsequently classify new ECG data as pre-ictal or normal.

The results show the validity of our proposal. In particular, the experiments demonstrate that HRV features are sufficiently discriminative. However, they also show that a SVM trained with data from one patient is almost useless with a different patient. Thus, the system must be properly trained for the specific patient who is going to use it.

Acknowledgement

This work is partially supported by Ministry of Economy and Competitiveness under contracts TIN2014-58662-R (AEI,FEDER,UE) and DPI2017-86372-C3-3-R (AEI,FEDER,UE).

REFERENCES

Al-Aweel et al., 1999. *Postictal heart rate oscillations in partial epilepsy. Neurology*, 53, no. 7, 1590–1592.

Baig, M.M. et al., 2013. *A comprehensive survey of wearable and wireless ECG monitoring systems for older adults. Medical and Biological Engineering and Computing*, 51, no. 5, 485–495.

Burguera A., 2018. *Fast QRS detection and ECG compression based on signal structural analysis. IEEE Journal of Biomedical and Health Informatics*, 23, no. 1, 123–131.

Eadie M., 2012. *Shortcomings in the current treatment of epilepsy. Expert Review of Neurotherapeutics*, 12, no. 12, 1419–1427.

Elgendi, M. et al., 2014. *Revisiting QRS detection methodologies for portable, wearable, battery-operated, and wireless ECG systems. PLoS ONE*, 9, no. 1.

Fujiwara, K. et al., 2015. *Epileptic Seizure Prediction Based on Multivariate Statistical Process Control of Heart Rate Variability Features. IEEE Transactions on Biomedical Engineering*, 9294, no. c, 1–1.

Gadhoumi, K. et al, 2016. *Seizure prediction for therapeutic devices: A review. Journal of Neuroscience Methods*, 260, 270–282.

Goldberger, A.L. et al., 2000. *PhysioBank, PhysioToolkit, and PhysioNet Components of a New Research Resource for Complex Physiologic Signals. Circulation*, 101, no. 23, e215–e220.

Hashimoto, H. et al., 2013. *Heart rate variability features for epilepsy seizure prediction.* In *2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA 2013.*

Pandit, D. et al., 2017. *A lightweight QRS detector for single lead ECG signals using a max-min difference algorithm. Computer Methods and Programs in Biomedicine*, 144, 61–75.

MODEL BASED EVALUATION OF INTEGRATED CARE IN HEART FAILURE TREATMENT

Alexander Lassnig, Theresa Rienmüller, Christian Baumgartner and Jörg Schröttner
Institute of Health Care Engineering
Graz University of Technology
A-8010, Graz,
Austria
E-mail: alexander.lassnig@tugraz.at

KEYWORDS
Agent Based, Discrete Event, Heart Failure Model, Hybrid Modeling, Telemonitoring.

ABSTRACT

Demographic changes, increased life expectancy and the associated rise in chronic diseases pose challenges to public health care systems. Optimized treatment methods and integrated concepts of care are potential solutions to overcome increasing financial burdens and improve overall quality of care. In this work a novel heart failure treatment model based on agent based and discrete event methodologies is presented. The hybrid approach facilitates the incorporation of different therapeutic procedures for outpatient and inpatient care on a patient individual level. Simulation results for conventional care and the use case of an exemplary telemonitoring program give deeper insight into cost distribution between outpatient and inpatient care and showcase potential positive health and economic outcomes within two years for integrated care. The presented model allows performing extensive simulations of established treatment procedures for heart failure patients and evaluating new holistic methods of care and innovative study designs. This approach offers health care providers a unique, adaptable and comprehensive tool for decision making in the complex and socioeconomically challenging field of cardiovascular diseases.

INTRODUCTION

Trends in demographics, emphasizing the population gap between young and old, increased life-expectancy and the associated rise in chronic diseases challenge public health care systems. Heart failure (HF), as the leading cause of hospitalizations of people above the age of 65 years is one example of cardiovascular diseases requiring explicit attention. (Zannad et al. 2009) HF-related treatment expenditures in high-income countries already encompass 2-3 % of the total health care budget and are projected to increase by more than 200 % in the next 20 years. (Heidenreich et al. 2013) In addition to the financial impact, HF is associated with an unfavorable prognosis. High mortality of up to 50 % within five years after the initial diagnosis underlines the severity of the disease. (Lloyd-Jones et al. 2010) One year case fatality after hospitalized HF is up to 30 % and disease-related admission rates are over 23 % in contrast to 12.6 % for all-cause readmissions, indicating room for improvement in post-inpatient management. (Gruneir et al. 2018) Patients' poor adherence to medication and recognition of early signs of cardiac decompensation, as well as insufficient collaboration among health care providers, are exemplary limitations in therapy. (Heidt et al. 2014) Estimations indicate that up to two thirds of HF readmissions are triggered by potentially preventable factors, including suboptimal discharge planning, non-adherence to HF medication, inadequate follow-up, insufficient social support as well as delays in seeking medical attention. New solutions based on optimized and individualized treatment and integrated concepts of care are potential ways to manage future burdens by detecting symptoms at an earlier stage and thus stabilizing the patient's health status and minimizing unnecessary admissions. (Koehler et al. 2011)
In this context, modeling is a powerful tool to evaluate potential benefits of treatment procedures on health and financial outcomes. This work presents the application of a novel hybrid modeling approach to simulate different treatment procedures for HF based on extensive data sets from outpatient and inpatient care to lastly support decision making in healthcare.

METHODS

The presented model builds upon a previously published HF treatment model (Schroettner and Lassnig 2013, Lassnig et al. 2017), further advancing the modeling methodology. The original discrete event model was complemented by an agent based approach, combining advantages of both methodologies. Discrete steps, implemented with their respective transition probabilities directly match the flow chart nature of the clinical pathway of HF patients. The agent based approach allows including individual patients with distinct features. Each entity in the model is represented as an agent of the class 'Patient'. Parameters such as age, gender, state of health and the history through the course of treatment classify each individual and influence transition probabilities along the decision tree of the discrete model (see Figure 1). The state of health is expressed by the functional classification scheme of the New York Heart Association (NYHA), describing four states of impairment, correlating with different frequencies, lengths and costs of the treatment procedures. (Yap et al. 2015) Class I would be a cardiac disease without the resulting limitations of physical activity, whereas class IV patients are unable to carry on any physical activity without severe discomfort. Every simulation run begins at the patient pool, where the cohort to be simulated is generated. The next step is to face transition probabilities for admissions, the outpatient clinic, the physician or potential death. In case of an admission to the hospital there may be the necessity for a stay at an intensive

care unit (ICU) or intermediate care unit (IMC). Patients can then potentially change their state of health, which leads to either further admissions or death. During admissions individual medical procedures (IMP) may also be necessary to improve health. After hospital discharge, in case there is no death event, patients are fed back to the patient pool. For outpatient clinics, physicians and specialists the same routines apply with the addition of medication.

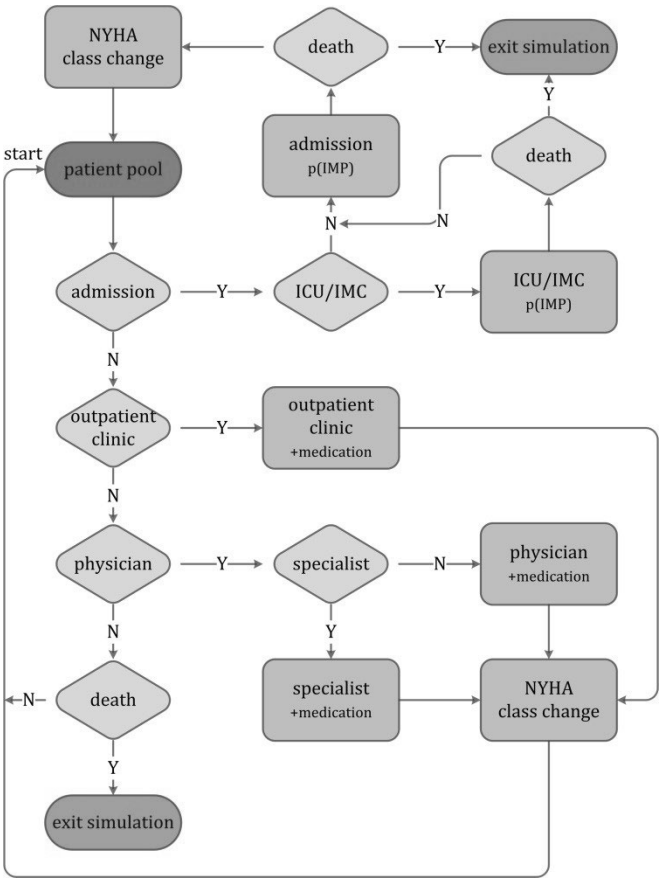


Figure 1: Flow Chart of the Clinical Pathway Implemented in the Model.

For simulation, the Java based software AnyLogic® (Version 8.3) was used. Overall clinical data on over 7,412 admitted HF patients and 25,939 visits to outpatient clinics in Austria served as the basis for the model construction, with additional data from an Austrian health insurance provider on 10,775 patients with information on outpatient medication. Details on implementations of each individual element of the flow chart can be taken from Lassnig et al. 2017. For this work the data basis could be expanded and analyses of individual patient reports by the health care provider were included, strongly improving NYHA coverage and the realistic simulation of HF related treatment. As a major improvement, individual medical procedures were fully addressed in terms of type, frequency and costs for inpatient care. The model flow in general was revised and all state transitions expressed with probability distribution functions depending on age, gender and the NYHA class.

Discrete Model

The discrete model is described as a hidden Markov model with distinct states q_i and transition probabilities p_{ij} (see Figure 2). The transition probabilities are derived from rate constants per day (sampling size $\Delta t = 1$) taken from the data sets. Discrete states are the inactive state in the patient pool, the physician, the specialist, the outpatient clinic and the hospital. For the discrete system the probability P_i of being in state q_i at time $k+1$ can be derived from the probability P_i at time step k and the outgoing and incoming probabilities of state q_i in the following way: (see Equation 1)

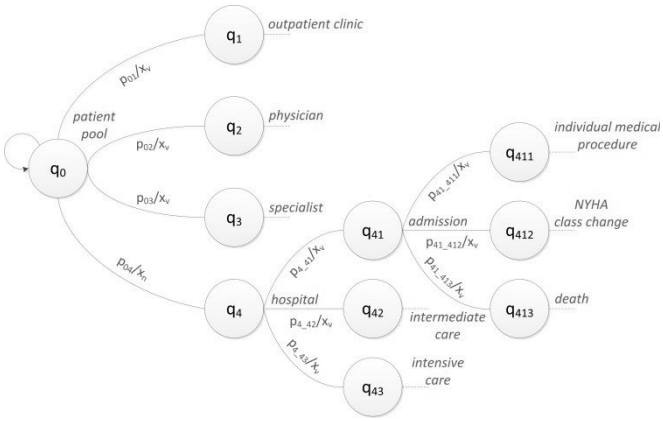


Figure 2: Exemplary State Transitions in the Hybrid Model, Starting with the Ground State q_0 of Patients in the Patient Pool. x_v are the Inner States of the Patient and p_{ij} the Transition Probabilities from State q_i to State q_j

$$P_{i,k+1} = \left(\sum_{j=1}^{N} P_{j,k} p_{ji} - \sum_{j=1}^{N} P_{i,k} p_{ij} \right) \cdot \Delta t + P_{i,k} \qquad (1)$$

where N is the total number of discrete states, p_{ij} describes the conditional probability of finding the system in a new state q_j, if it has recently been in state q_i. (p_{ij} corresponds to transitions out of state q_i and p_{ji} to transitions entering state q_i).

Extended Hybrid Model

In the presented hybrid model the discrete model is combined with an agent based approach. For the simulated scenarios in the Results section the probabilities for state transitions p_{ij} depended on the following inner states $\vec{x}_v$ of the agent v:

$$\overrightarrow{x_v} = \begin{bmatrix} \text{age} \\ \text{sex} \\ \text{NYHA} \end{bmatrix}, v=1,\ldots,n \qquad (2)$$

Based on comprehensive data these three inner states can be further expanded to e.g. also investigate effects of comorbidities or risk factors. Additionally each agent also contains a set of auxiliary variables, logging necessary information per agent on the course of treatment. These variables comprise e.g. costs, frequencies of visits, lengths of stay (LOS).

The initial inner states of the patients follow set values or probability distributions that can be defined at the beginning of the stimulation. In our use cases these probability distributions were defined mainly based on data sets of Austrian health insurance and health care providers. After leaving a specific state q_i there are potential changes of the inner state x_v of the patient which in turn may alter the state

transition probabilities. Investigating selective problems and scientific questions can be carried out by changing state transition probabilities. In case of death, patients exit the simulation run and do not interact with the virtual environment anymore.

RESULTS

In this work a novel heart failure model is presented that introduces an agent based modeling approach to a discrete event setting, focusing on the impact of the state of health on economic and health outcome. The basis for the simulations was the following: 10,000 patients (equal gender and age distributions between 50 and 85 years of age) were simulated over a time span of 5 years, with 2,500 patients in each of the four NYHA classes. Individual simulation runs are based on a random number generator initialized with random seeds. 100 simulation runs were compared for each scenario in order to narrow statistical deviations and improve consistency of results, which were expressed with mean values plus standard deviations in the figures.

Scenario 1 – Conventional Care

Disregarding NYHA class changes and mortality rates, Table 1 sums up overall costs per patient and year.

Table 1: Mean Annual Costs plus standard deviation (SD) per Patient and Year for Each NYHA Class (in €)

	NYHA class			
	I	II	III	IV
annual costs (SD)	2,018 (14)	3,639 (14)	5,459 (58)	10,077 (165)

Treatment efforts for NYHA class IV patients, with an average of €10,077 ± €165, were significantly higher than corresponding values of other classes, mostly due to longer and more frequent admissions. Figure 3 shows a breakdown of costs between outpatient (OP) and inpatient (IP) care per NYHA class.

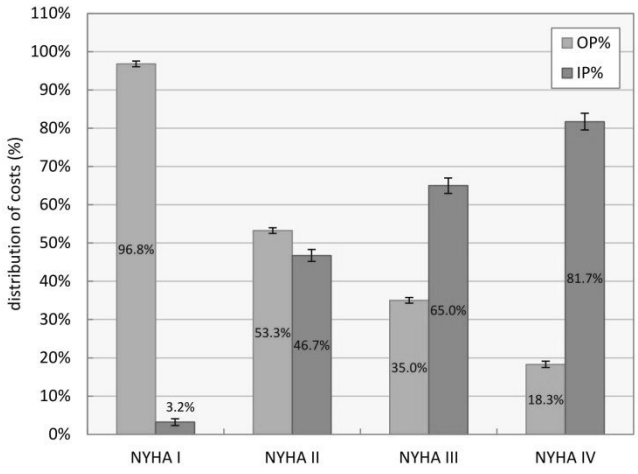


Figure 3: Costs Distribution between Outpatient (OP) and Inpatient (IP) Care for the Four NYHA Classes

The distribution of treatment efforts in outpatient and inpatient care correlates with the severity of the heart failure condition. In the data set, NYHA class I patients were rarely treated in inpatient care, while for NYHA class IV patients; admissions amounted to over 80 % of the related expenses. Overall costs for outpatient care were divided into expenses (median values) for the physician (27.6 %), the outpatient clinic (9.4 %) and the prescribed medication (63.0 %).

Scenario 2 – Telemonitoring Program

To simulate the use case of a telemonitoring program, two different applications have been tested. The additional costs of the telemonitoring approach were calculated to €1,000 per patient as initial expenses for equipment purchase and an extra €45 per patient and month to cover additional efforts by the physician and for maintenance and service of the system. The impact of additional efforts in outpatient care was extracted from Dendale et al. 2012 and amounted to 21 % decrease in overall admission rate, as well as a reduction of the average length of stay by 35 %. In Figure 4 expenses for conventional care are compared with the telemonitoring systems TM_1 (as described above) and TM_2 (no telemonitoring system for NYHA class I patients).

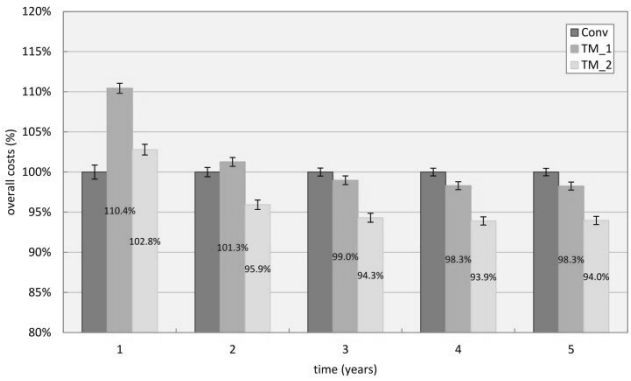


Figure 4: Overall Costs for Conventional Care (Conv) and Two Different Implementations of the Same Telemonitoring Program

Initial investments for the chosen telemonitoring system increased costs within the first year by up to 10 %. Depending on the telemonitoring approach, cost efficiency could be reached within the first two years, after three years both systems were cost efficient compared to conventional care.

DISCUSSION

There are several decision-analytic modeling approaches reported in literature, estimating effects of health technologies for chronic HF patients. Most are focusing on the effectiveness of new pharmacological or device-oriented interventions. (Goehler et al. 2011) A comprehensive analysis of survival in HF treatment has been published by Levy et al. 2006 with the Seattle Heart Failure Model, outcomes for overall admission and death were investigated by Gasperoni et al. in 2017. In contrast to published models, the detailed description of inpatient and outpatient care as delineated in this work in regard to the NYHA classification system is unique and has high potential for future application.

The presented simulation results give an overview of exemplary model capabilities for conventional care and a telemonitoring program. The clear shift of costs with worsening heart condition (Table 1) and the resulting distribution between outpatient and inpatient care corresponds to clinical guidelines and findings in literature. In a systematic review of the economic burden caused by HF, Shafie et al. in 2018 reported huge gaps of costs in literature, where annual expenses for the treatment of NYHA class IV patients ranged from Int$4,147 to Int$36,297 and from Int$3,604 to Int$20,871 for NYHA class III. Median annual inpatient costs per person summed up to Int$10,141. The simulated outcome of €10,077 ± €165 per NYHA class IV patient and year correlates with these findings.

The presented use case of a telemonitoring system, based on findings by Dendale et al. in 2012, scores in the upper spectrum of potential benefits, with a decrease in admission rates as well as length of stay. Primary investments for the implementation of the system increase costs at first, but can turn out to be cost efficient within 2 years, as depicted in Figure 4. These findings are highly sensitive to the chosen patient collective in terms of age and overall state of health. Results indicate that the highest potential for the application of a telemonitoring system can be achieved for patients in the NYHA classes II and III, which are more prone to potential changes of the state of health.

With the herein reported model, a variety of approaches can be modeled to estimate outcomes for highly specific patient collectives and to directly assess potential solutions. With the economic outcomes being within ~1 % of the median of the published review by Shafie et al. in 2018, a first approach to validate simulation results has been taken. In general reported costs strongly vary in literature and are heavily dependent on the individual health care system and treatment setting, requiring careful interpretation and estimations.

The model can only be a reflection of the real world and several assumptions have been made to draw conclusions. The lack of information on the cause of death disallowed the simulation with outpatient mortalities based on HF. For the presented results, comorbidities haven been neglected as well as the influence of ethnicity and social status. Data restriction policies of the health care and health insurance provider disallowed the conjunction of the data sets. Access to consistent and comprehensible patient profiles across all areas of care could significantly improve model performance.

CONCLUSION

Simulation results show that basic outcomes for conventional care comply with findings in literature and give further insight into cost distribution based on the severity of HF. The exemplary use of a telemonitoring program revealed potential cost savings within the time frame of two years even without regarding positive effects on the state of health. Comprehensive simulations of established treatment procedures provide the basis for the evaluation of new holistic methods of care and innovative study designs to improve health and economic outcomes. This work offers health care providers a novel tool for decision making in the complex and socioeconomic challenging field of cardiovascular disease and thus may be a further piece in the sophisticated puzzle of refining health care systems

REFERENCES

Berry, C., Murdoch, D., and McMurray, J. 2001. “Economics of chronic heart failure”. *European Journal of Heart Failure*, No.3, 283-291.

Dendale, P.; de Keulenaer, G.; Troisfontaines, P.; Weytjens, C.; Mullens, W.; Elegeert, I.; Ector, B.; Houbrechts, M.; Willekens, K.; Hansen, D.; 2012. “Effect of a telemonitoring-facilitated collaboration between general practitioner and heart failure clinic on mortality and rehospitalization rates in severe heart failure: the TEMA-HF 1 (TElemonitoring in the MAnagement of Heart Failure) study”. *Eur. J. Heart. Fail.* 14, 333–340.

Gasperoni F, Ieva F, Barbati G, Scagnetto A, Iorio A, Sinagra G, Di Lenarda A. Multi-state modelling of heart failure care path: A population-based investigation from Italy. *PLoS ONE*. 2017;12:e0179176.

Goehler A, Geisler BP, Manne JM, Jahn B, Conrads-Frank A, Schnell-Inderst P, et al. Decision-analytic models to simulate health outcomes and costs in heart failure: a systematic review. *Pharmacoeconomics*. 2011;29:753–69.

Gruneir, A.; Fung, K.; Fischer, H.D.; Bronskill, S.E.; Panjwani, D.; Bell, C.M.; Dhalla, I.; Rochon, P.A.; Anderson, G. 2018. „Care setting and 30-day hospital readmissions among older adults: a population-based cohort study“, *CMAJ* 190, E1124-33.

Koehler, F., Winkler, S., Schieber, M., Sechtem, U., Stangl, K., Böhm, M., et al. 2011. „Impact of Remote Telemedical Management on Mortality and Hospitalizations in Ambulatory Patients With Chronic Heart Failure”. *Circulation*, 123, 1873-1880.

Heidenreich, P.A.; Albert, N.M.; Allen, L.A.; Bluemke, D.A.; Butler, J.; Fonarow, G.C.; Ikonomidis, J.S.; Khavjou, O.; Konstam, M.A.; Maddox, T.M.; Nichol, G; Pham, M.; Piña, I.L.; Trogdon, J.G. 2013. “Forecasting the impact of heart failure in the United States: a policy statement from the American Heart Association”. *Circ. Heart. Fail.* 6, 606–619.

Heidt, A.; Ammenwerth, E.; Bauer, K.; Fetz, B.; Fluckinger, T.; Gassner, A.; Grander, W.; Gritsch, W.M.; Haffner, I.; Henle-Talirz, G.; Hoschek, S.; Huter, S.; Kastner, P.; Krestan, S.; Kufner, P.; Modre-Osprian, R.; Noebl, J.; Radi, M.; Raffeiner, C.; Welte, S.; Wiseman, A.; Poelzl, G. 2014. “HerzMobil Tirol network: rationale for and design of a collaborative heart failure disease management program in Austria”. *Klin. Wochenschr.* 126, 734–741.

Lassnig, A.; Baumgartner, C.; Schroettner, J. 2017. “A modeling approach to heart failure treatment”. ESM’2017, 425-427.

Lloyd-Jones, D.; Adams, R.; and Carnethon, M. 2009. “Heart disease and stroke statistics-2009 update: A report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee”. *American Heart Association*, No.119, e21-e181.

Ross, J.S.; Chen, J.; Lin, Z.; Bueno, H.; Curtis, J.P.; Keenan, P.S.; Normand, S.-L.T.; Schreiner, G.; Spertus, J.A.; Vidán, M.T.; Wang, Y. ; Wang, Y.; Krumholz, H.M. 2010. “Recent national trends in readmission rates after heart failure hospitalization”. *Circulation*, No.3, 97-103.

Schroettner, J. and Lassnig, A. 2013. “Simulation model for cost estimation of integrated care concepts of heart failure patients”. *Health Economics Review*, 3:26.

Shafie, A.A.; Tan, Y.P.; Ng, C.H.; 2018. “Systematic review of economic burden of heart failure”. *Heart. Fail. Rev.* 23, 131–145.

Yap, J.; Lim, F.Y.; Gao, F.; Teo, L.L.; Lam, C.S.P.; Yeo, K.K. 2015. “Correlation of the New York Heart Association Classification and the 6-Minute Walk Distance: A Systematic Review”. *Clinical cardiology 38*,621–628.

Zannad, F.; Agrinier, N. and Alla, F. 2009. “Heart failure burden and therapy”. *Europace*, No.11 (Suppl. 5), 1-9.

MODELING AND OPTIMIZATION OF COMBINED CHEMO-RADIOTHERAPY

Andrzej Swierniak
Jaroslaw Smieja
Krzysztof Fujarewicz
Department of Automatic Control, Silesian University of Technology, Akademicka 16,
44-101 Gliwice, Poland,
E-mail:andrzej.swierniak@polsl.pl

Rafal Suwinski
M. Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Gliwice Branch.
Wybrzeze Armii Krajowej 15,
44 101 Gliwice, Poland
E-mail:rafal.suwinski@io.gliwice.pl

KEYWORDS
Biomathematical modeling, optimization, cancer

ABSTRACT

We summarize results of our research studies on models of combined anticancer therapy and their comparison with real clinical data. We use two mathematical techniques which, to our knowlegde, have not been applied simultaneously: optimal control theory and survival analysis. Optimal treatment protocols are found using Gompertz model of tumor growth. Then both structural and parametric sensitivity is analysed. Finally we present results of survival analysis based on the Kaplan-Meier curves for different protocols of chemo-radio-therapy and compare them with real clinical data and results of optimal treatment protocols.

INTRODUCTION

Pre- or post-operational chemo-radio-therapy have become one of the standard adjuvant therapies in recent years. Though both chemo- and radiotherapy protocols are standardized, the question about the order in which they should be applied, or concurrency, remains an open question. In this work we attempt to answer it with mathematical modeling and optimization of two-dimensional control that represents therapy in a control-theory based approach (see e,g, Swierniak et al. 2016, Schaettler and Ledzewicz 2015 and references therein). Practical application of such theoretical work, however, are scarce. Nevertheless, recent progress in linking various biomolecular markers to the stage of the disease as well as biomedical imaging has provided methods of monitoring cancer growth, studying cancer cell populations and that creates a new opening for mathematical modeling in this field and refining control theory methods so that they might gain acknowledgment in clinics.). In order to address this problem, two issues are discussed. First, two different ways of modeling tumor growth under therapy are compared. For each of them the necessary conditions for optimal control representing the therapy are presented and discussed. Then, Kaplan-Meier survival curves are compared for standard therapy protocols used in clinics and different approaches to model tumor growth. Despite the fact that the literature devoted to optimization of anticancer treatment is large, in general, combination of radio- and chemotherapy has escaped attention of the researchers. There have been only few efforts to model them and to compare concurrent versus sequential chemo-radio-therapy (e.g. Geng et al., 2017, Curran et al. 2011). However, they were focused on testing specific protocols, and no attempt to optimize therapy has been made there. On the other hand, when treatment is dealt with as a control optimization problem, it is not radio- and chemo-therapy combined. Moreover, in optimization-oriented publications the models do not relate to survival curves which arguably constitute the only point of interest that could gain the attention of clinicians. The goal of this paper is to answer two questions, important from the point of view of clinical relevance of modeling and optimization of anticancer treatment: (i) how the necessary conditions derived for an control optimization problem, representing the search for the best therapy protocol, depend on the model structure? and (ii) is the choice of the cancer growth model important for survival curves used to evaluate the ultimate outcome of the therapy? In that respect, we analyze structural sensitivity of control models involving therapy strategies with respect to different tumor growth terms and different ways of modeling the therapies. In this work we combine two approaches that usually do not meet in mathematical modeling of anticancer treatment. First, after introducing the model of cancer growth and treatment, optimal control problem is stated. Necessary conditions are provided and analyzed in the context of their clinical relevance. Then, Kaplan-Meier survival curves (Dudley et al. 2016) for patients with cancer are shown for a particular numerical example and the question of their dependence on a particular model form is discussed.

OPTIMIZATION OF TREATMENT

Let us assume that the population of cancer cells, described by its size N(t), when no treatment is administered, grows according to a Gompertz model, one of the most frequently used in modeling population dynamics. Let us introduce two-dimensional control vector [u(t),d(t)], whose components represent chemotherapy and radiotherapy, accordingly. The chemotherapy usually is assumed to kill the tumor cells according to the log-kill hypothesis (Skipper et al. 1964), hence the kill rate is proportional to chemotherapeutic agent concentration at time t. Radiotherapy effect is usually described in terms of the so called linear-quadratic (LQ)

model (Fowler 1989). The goal of the therapy is formulated as a problem of maximization of tumor control probability (TCP). Assuming constant values of tumor cells density and clonogenic fraction it can be transformed to the problem of finding measurable functions u and d that minimize tumor volume at the end of treatment, subject to the constraints imposed on the maximal and cumulated doses of drugs and radiation. If an oversimplified case is assumed, in which a chemotoxic drug instantly affects cancer cells, following its administration (no pharmacokinetics considered) and no DNA repair is taken into account, application of the Pontryagin's maximum principle leads to the bang- bang (switching) form of optimal control in which the no treatment interval is followed by the maximal treatment dose. This means that, somehow counterintuitively, both radiation and chemotherapy should be applied after a delay. Moreover, they will always be applied simultaneously over a final part of treatment. However, this model ignores two treatment-related phenomena drug metabolism of chemotherapeutic agent (pharmacokinetics – PK) and DNA repair of adjacent strand breaks. As they may affect the structure of theoretically optimal treatment protocols, we propose to incorporate their mathematical description into the model. An important finding here is that while the optimal protocol of radiotherapy has been found to be the same regardless of incorporating or neglecting DNA repair, the form of optimal chemotherapy is affected by including PK in the system description. When PK is explicitly included in the model, the optimal solution suggests two possible strategies, where the exact choice depends on model parameters. In one case, the optimal strategy is completely opposite to the solution obtained without PK, i.e. it suggests applying chemotherapy at the beginning of the treatment. In the other case, the optimal strategy involves postponing chemotherapy administration, but not quite to the very end of treatment horizon. This may be attributed to the fact that the treatment effect persists even after drug administration due to PK. The above finding gives rise to another question, about the implications of assuming only one of possible descriptions of tumor growth.

SENSITIVITY OF THE MODEL

As shown in the previous section, expanding mathematical description of the system under investigation by incorporating additional processes may lead to qualitatively different results, as far as optimal control is concerned. The interesting finding is that although the structure of optimal control changes the results of that are almost negligible, as the resulting dynamics of tumor growth is considered. On the other hand, some processes, for example tumor growth, may be described by different models. Moreover, even if the optimal control law is found to be always in the bang-bang form, the switching times depend on model parameters. That means that they may be different for individual patients. To check if the choice of growth model affects therapy outcome, instead of deriving necessary conditions for each case separately, we focus on therapy outcomes for an a priori assumed treatment protocols and different models of tumor growth. As an example of the similar but having different properties we consider a logistic growth. We also have studied results for an exponential growth model. The main difference between the Gompertz and the logistic models is that the first one assumes an exponentially decreasing growth rate per capita, while the second one assumes linearly decreasing growth rate per capita. One of the important implications of this difference is that in the case of Gompertz model, a complete eradication of the „virtual" tumor is impossible The results from the previous section indicated that a bang-bang solution is optimal for a given fixed terminal time. Moreover, current clinical protocols involve multiple bang-bang controls that may be interpreted as subsequent cycles of therapy, with optimal control applied in each cycle. Therefore, three different multivariable control strategies have been checked: sequential chemo- and radiotherapy (CRT), concurrent CRT and radiotherapy only. Realistic clinical protocols were tested, where in each case the therapy was started 14 days after tumor detection. Recently, parameter values for the presented model (for Gompertz model of tumor growth) were estimated, based on survival curves available in literature (Gang et al. 2017). Additionally, we applied the following assumptions: 1) the patient dies when tumor reaches 13 [cm] in diameter (the death condition), 2) there is a correlation between the tumor growth rate and the radiosensitivity parameter α (coefficient = 0.87, estimated based on in vitro experiments, Lee et al. 2016), 3) the patient is cured after the treatment if random number generated from uniform distribution [0,1] is lower than the tumor control probability.

Three treatment strategies were tested on a group of a thousand virtual patients. Each patient was characterized by different parameter values, drawn from bivariate normal distribution to reproduce high correlation between parameters. Firstly, we compared results obtained by three models for a group of thousand patients grouped by treatment strategy. Results are presented in the form of Kaplan-Meier survival curves . There is a significant difference between treatment outcome for concurrent CRT and RT only, for all models. When concurrent CRT therapy was applied the lower overall survival was predicted by Gompertz model. The differences in Kaplan-Meier curves are negligible in short time after the treatment, but they increase in time. Using Gompertz and logistic model, it was predicted that 5 years after tumor detection, respectively, 11% and 18% of patients would stay alive. When only radiotherapy was applied the higher overall survival was predicted with the Gompertz growth model. The order is different than previously (when concurrent CRT was applied). Additionally, in this case, the differences between Kaplan-Meier curves are visible during the whole period of

observation. The biggest differences between Gompertz and logistic models occur one year after tumor was detected, and overall survival at this time is 81% and 69%, respectively. Secondly, we analyzed the effect of different treatment strategies on each model. Concurrent CRT therapy gave the highest overall survival after 5 years for all analyzed models, where only RT gave the lowest one. Sequential and concurrent CRT applied for Gompertz model gave the same treatment outcome, which suggests that this model is less sensitive to control. Results obtained by logistic models suggest that application of sequential or concurrent CRT may have significant impact on overall survival of patients. There are some differences between overall survival of patients obtained by three different models. However, the shape of curves were similar and it may be related to tumor growth rate parameter value. With appropriate estimation of parameter values any of presented models could reproduce clinical data with similar accuracy. In all cases, the most effective treatment was concurrent chemo-radio-therapy. Exponential and logistic models were more sensitive to changes of therapy protocols than the Gompertz model. The Kaplan-Meier curves for the best case has been compared to the ones for the optimal protocols found in the previous session. The interesting finding is that surprisingly the differences between those curves are not significant.

DISCUSSION

Tumor growth and its treatment can be viewed and analyzed as an open–loop control system. Thus, optimal treatment protocols can be found using formalisms of optimization theory. The particular form of optimal control in terms of the order in which chemo and radiotherapy should be applied depends on including or neglecting pharmacokinetics of chemotoxic drugs. However, in both cases the necessary conditions suggest that both therapies should be applied concurrently, at least in some period of time. Incorporating or neglecting DNA repair mechanisms do not have the same impact on the form of necessary conditions (in both cases the solution is qualitatively the same). Furthermore, survival analysis, in which the actual clinical protocols for chemo- and radio-therapy were applied, confirms the above conclusion, showing that concurrent therapies yield significantly better survival prospects. This conclusion does not depend on the type of tumor growth that is used in the model. One way to extend the ideas studied above is to use a hybrid model in which parameters of the tumor growth model in the presence of treatment are adjusted using clinical data and the criterion of adjustment is based on the differences between real and model survival curves (Wolkowicz et al, 2019) Yet another idea is to incorporate synergistic effects of the two modalities (radio and chemo-therapies) in the model of tumor growth (see e.g. Bajger et al., 2019).

ACKNOWLEDGEMENT

The Authors would like to thank for financial support of their research. The study is partially supported by National Science Committee, Poland, Grant no.2016/21/B/ST7/02241 and partially by Silesian University of Technology Grant no.02/010/BK18/0102.

REFERENCES

Bajger, P., Fujarewicz, K., Swierniak, A.: Effects of pharmacokinetics and DNA repair on the structure of optimal controls in a simple model of radio-chemotherapy. In: *Proc of the MMAR Conference*, pp. 686-691, (2018)

Bajger P., Fujarewicz K., Swierniak A.: Optimal control in a model of chemotherapy-induced radiosensilization, *Mathematica Applicanda*, 28/29, 7-18, (2019).

Curran, W.J. Jr, Paulus, R., Langer, C.J., Komaki, R., Lee, J.S., Hauser, S., Movsas, B., Wasserman, T., Rosenthal, S.A., Gore, E., Machtay, M., Sause, W., Cox, J. D.: Sequential vs concurrent chemoradiation for stage III nonsmall cell lung cancer: Randomized phase III trial RTOG 9410. JNCI: Journal of the National Cancer Institute, 103(19), 1452-1460, (2011).

Dolbniak M., Smieja, J., Swierniak, A.: Structural sensitivity of control models arising in combined chemo-radiotherapy. In: *Proc of the MMAR Conference,* pp. 339–344, (2018).

Dudley, W.N., Wickham, R., Coombs, N.: An Introduction to Survival Statistics: Kaplan-Meier Analysis. *J Adv Pract Oncol*, 7(1), 91–100 (2016)

Fowler,J. F.: The linear-quadratic formula and progress in fractionated radiotherapy. *Br J Radiol*. 62, 679-94 (1989).

Geng, C., Paganetti, H., Grassberger, C.: Prediction of treatment response for combined chemo- and radiation therapy for non-small cell lung cancer patients using a biomathematical model. *Sci Rep.* 7, (2017). DOI: 10.1038/s41598-017-13646-z.

Lee, J.Y., Kim, M-S., Kim, E. H. , Chung, N., Jeong, Y. K., Retrospective growth kinetics and radiosensitivity analysis of various human xenograft models, *Lab Anim Res*; 32, 187-193, (2016).

Schaettler, H., Ledzewicz, U.: *Optimal control for mathematical models of cancer therapies. An application of geometric methods*, Springer Science + Buisness Media, New York 2015

Skipper, H.E., Schabel, F., Wilcox, W.: Experimental evaluation of potential anticancer agents. xiii. on the criteria and kinetics associated with curability of experimental leukemia. *Cancer Chemother Rep* 35, 1-111 (1964).

Swierniak, A., Kimmel, M., Smieja, J., Puszynski, K., Psiuk–Maksymowicz, K.*: System engineering approach to planning anticancer therapies*, Springer International Publishing, Switzerland 2016

Wolkowicz S., Fujarewicz K., Kurpas M., Lakomiec K., Smieja J., Suwinski R., Swierniak A.: Prediction of lung cancer patients' response to combined chemo-radiotherapy using a personalized hybryd model, *Mathematica Applicanda,* 28/29, 141-150, (2019).

ON USING EEG SIGNALS FOR EMOTION MODELING AND BIOMETRICS

Miguel Arevalillo-Herráez, Guillermo Chicote-Huete, Francesc J. Ferri
Departament d'Informàtica. Universitat de València. Valencia, Spain
{miguel.arevalillo,guillermo.chicote,francesc.ferri}@uv.es

Aladdin Ayesh
Faculty of Technology. De Montfort University
Leicester, United Kingdom
aayesh@dmu.ac.uk

Jesús G. Boticario
Departamento de Inteligencia Artificial. UNED.
Madrid, Spain
jgb@dia.uned.es

Stamos Katsigiannis, Naeem Ramzan
School of Computing. University of the West of Scotland. Paisley, United Kingdom
{stamos.katsigiannis,naeem.ramzan}@uws.ac.uk

Pablo Arnau-González
CECOTEC Innovaciones, S.L
pabloarnau@cecotec.es

KEYWORDS
EEG, Affect, Biometrics, Emotion modeling

ABSTRACT

A number of previous works have adopted a subject independent approach for recognizing emotions from Electroencephalography (EEG) signals, and attempted to build a global model by treating data from different subjects as if they belong to the same individual. In this paper we visually explore the data provided in four different standard datasets when using Power Spectral Density features, and show that the subject-dependent component in the EEG signal is far stronger than the emotion-related component. In addition, the session-dependency that is also found discourages the application of this type of features from EEG signals in a biometric context.

INTRODUCTION

The existing relation between emotions and electrical activities of the brain (Soleymani et al. 2014) has motivated an extensive research on detecting emotions from electroencephalogram (EEG) signals, sometimes in combination with other sources of information (Lu et al. 2015, Koelstra et al. 2012, Soleymani et al. 2012a). In this context, we can clearly differentiate between two distinct approaches. On the one hand, subject-dependent methods (Sohaib et al. 2013, Salmeron-Majadas et al. 2015, Ayesh et al. 2016) create a model per individual, using training information that only refers to that person. On the other hand, subject-independent techniques (Wang et al. 2014, Hadjidimitriou et al. 2015, Liu et al. 2010) ignore specific subject traits, and train a single global model that aims to be valid for all users. The advantage of the latter strives on the fact that more training data becomes easily available due to combining data from all individuals in the dataset. It also means that no previous information about a specific subject is required to start the prediction task.

In this paper, we analyse the suitability of these two approaches for emotion detection from EEG signals when using Power Spectral Density (PSD) features. To this end, we use t-Distributed Stochastic Neighbor Embedding (t-SNE) (Maaten and Hinton 2008), a dimensionality reduction method which is specially suited for the visualization of high-dimensional datasets. As a first contribution, a visual exploration of the data reveals that samples from the same individual are naturally grouped in clusters, and suggests that the subject-related component of the signal has a relatively higher importance than the emotion-related component. Such observation reinforces the arguments provided in Ayesh et al. (2014), regarding the need to take into consideration a person individuality at expressing emotions. As a second contribution, and despite that this particular layout seems to support the suitability of EEG signals for biometric applications, we discover a strong dependency from the session that also hinders practical applications in this context. As a natural consequence, this opens the door to mixed or multi-task approaches in which emotion, identity and session jointly contribute to uncover hidden patterns in the data.

DATA ANALYSIS

Datasets

The growing interest in affect recognition from EEG signals has motivated the development of dedicated pub-

Table 1: Summary of characteristics for the databases in the study

Database	Number of participants	Number of videos	Video content	Video duration	Recording device	Number of channels	Sampling frequency	Number of sessions
DEAP	32	40	Music videos	60 s	Biosemi Active II	32	512 Hz (downsampled to 256 Hz)	1
MANHOB-HCI	27	20	Excerpts from movies	34.9-117 s (avg=81s)	Biosemi Active II	32	512 Hz (downsampled to 256 Hz)	1
DREAMER	23	18	Music videos	65-393 s (avg=199s)	Emotive EPOC	14	128 Hz	1
SEED	15	15	Excerpts from movies	240 s	ESI NeuroScan	62 (32 used)	1000 Hz (downsampled to 200 Hz)	3

lic datasets. In between them, we find DEAP (Koelstra et al. 2012), MAHNOB-HCI (Soleymani et al. 2012b), DREAMER (Katsigiannis and Ramzan 2017), and SEED (Duan et al. 2013, Zheng and Lu 2015). These databases are composed of EEG recordings from subjects that were exposed to a series of video stimuli aimed at eliciting specific emotions. They all use Russell's two dimensional bipolar emotional model to label and represent emotions. The two dimensions are valence (positive or negative feeling) and activation/arousal (level of excitement) (Russell 1979), which are justified to account for the major proportion of variance in affect scales. In DEAP, MAHNOB-HCI and DREAMER, affective labels for valence and arousal were self-reported by the user using self-assessment manikins (SAM) (Morris 1995). As in previous works by other authors (Arnau-González et al. 2017, Koelstra et al. 2012, Liu et al. 2015, Petrantonakis and Hadjileontiadis 2010), the reported value was discretized into positive/negative using 5 (out of 10) as the threshold value. In SEED, recordings were instead annotated according to the expected emotional response, using three possible discrete values for valence: negative, neutral and positive.
Out of the four datasets described above, the only one that considers the concept of a session is SEED. In SEED, the experimental setting was repeated three times across different dates, recording the data for each session independently. A summary of the characteristics of the four datasets we have used in the paper is provided in Table 1.

Features

PSD features have extensively been used for identifying emotions from EEG signals (Arnau-González et al. 2017, Katsigiannis and Ramzan 2017, Soleymani et al. 2012b, Koelstra et al. 2012). We computed the PSD in each of the available channels, discarding the information from the channels in SEED that were not available in DEAP and MANHOB. For consistency reasons and to ease the future comparison of results, this was done as in the original publications (Koelstra et al. 2012, Soleymani et al. 2012a, Zheng and Lu 2015), downsampling the signals to 128 Hz in the case of MANHOB, and using Welch's method with a Hamming window of 1 *sec*, with 50% overlapping. The spectral power was averaged over the θ (4-8 Hz), $\tilde{\alpha}$ (8-10 Hz), α (8-12 Hz), β (12-30 Hz), and γ (30+ Hz) bands for all channels or electrodes. In addition, the difference between the spectral power of all the symmetrical pairs of electrodes on the right and left hemisphere in the same bands was also computed in order to measure the possible asymmetry in the brain activities due to emotional stimuli. This yielded a vector of 230 features per recording in DEAP, MANHOB and SEED (32 electrodes $\times$ 5 bands + 14 pairs $\times$ 5 bands); and 105 in DREAMER (14 electrodes $\times$ 5 bands + 7 pairs $\times$ 5 bands).

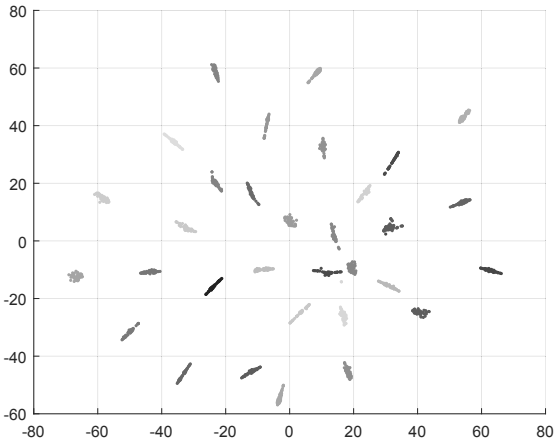
Dimensionality Reduction Experimentation

Experimental setting

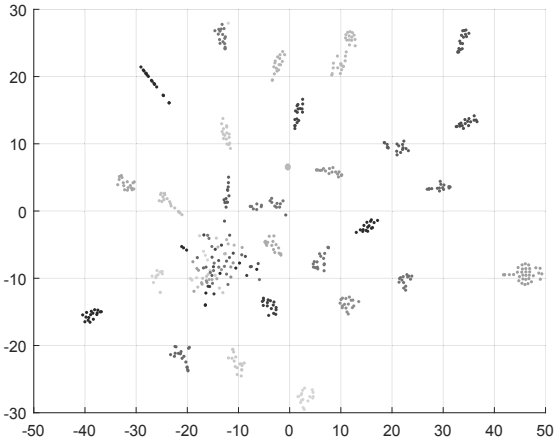
t-Distributed Stochastic Neighbor Embedding (t-SNE) (Maaten and Hinton 2008) is a technique for dimensionality reduction that is particularly well suited for the visualization of high-dimensional datasets. We have applied this method to EEG recordings from different individuals so as to reduce the dimensionality of these datasets into a two dimensional (2-D) space. In order to ensure consistency of the results, experiments have been repeated across the different databases described above.

Results and discussion

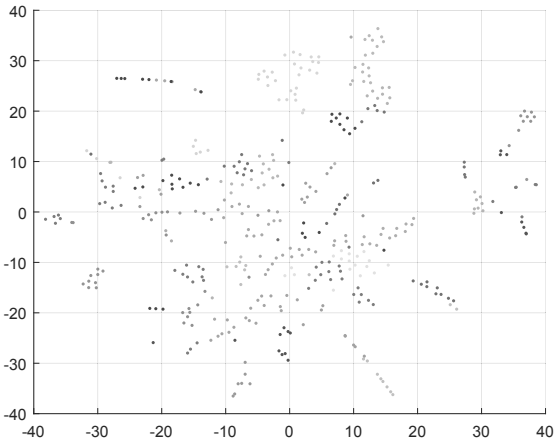
Figure 1 shows a 2-D representation of the data recordings stored in the DEAP, MANHOB and DREAMER datasets. The same kind of data for a representative session in SEED is shown in Figure 2.
In all these plots, we have used the same color to represent recordings that belong to a same subject. This allows one to easily observe that EEG recordings are



(a)



(b)



(c)

Figure 1: Dimensionality reduction by t-SNE in (a) DEAP, (b) MANHOB and (c) DREAMER.

naturally clustered according to the subject they are associated with. This clustered structure is more obvious in DEAP and MANHOB, but still noticeable in DREAMER and SEED.

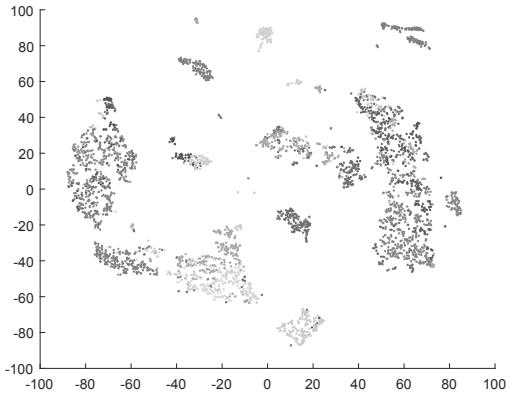
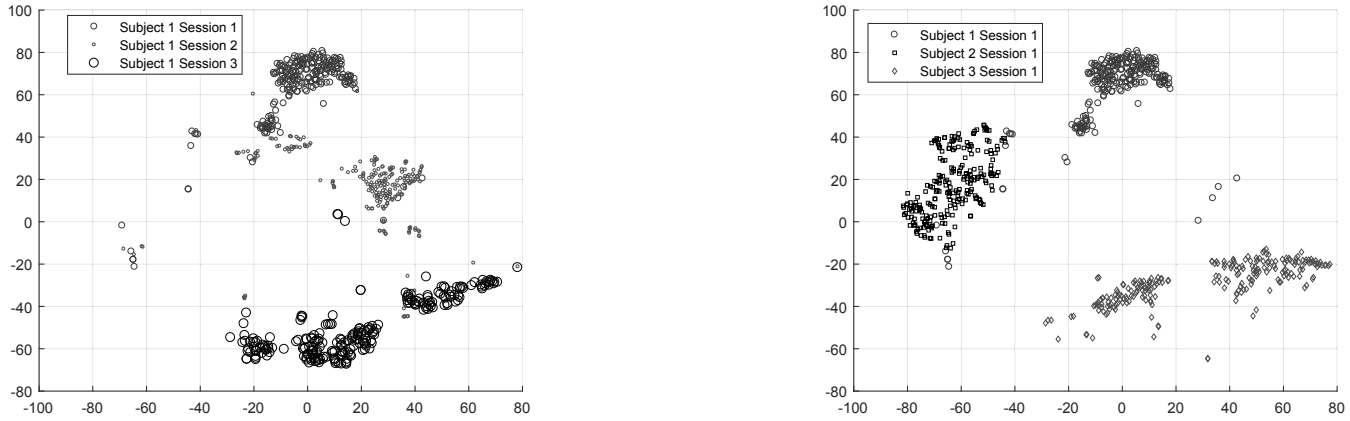


Figure 2: Dimensionality reduction by t-SNE for a session in SEED.

In all cases, samples (recordings) that belong to a same person appear close to each other, independent from the emotion elicited. This suggests that personality traits are of the factors that most influences PSD features from EEG signals, above others including the subject's affective state. Such result is against the construction of a subject-independent model that is valid for unseen users, and suggest that positive classification results obtained by some authors are more likely due to the class imbalance in the datasets used in their experiments than to the adequacy of the procedure for the problem at hand. For example, accuracy reported in (Katsigiannis and Ramzan 2017) is around 0.62 in valence and arousal, when using a SVM with a Radial Basis Function (RBF) kernel. However, the database was unbalanced (56%-44% in arousal and 61%-39% in valence).

Although the plots shown in Figures 1 and 2 clearly discourages the use of subject-independent modelling approaches, they also support the use of EEG signals for biometric authentication. However, one basic characteristic of suitable biometric authentication methods relate to permanence, i.e., the trait should be reasonably invariant over time with respect to the specific matching algorithm. One way to test permanence is using multiple session data. Figure 3 (a) shows an example of a 2-D data representation of three session for a single user in the SEED database. A strong dependency from the session can be observed in this plot, suggested by the clusters than can be easily distinguished. This indicates that there are aspects related to the session, such as the position of the electrodes, that have a considerable impact on the signal. This effect plays against the permanence property of EEG signals across multiple sessions and it is consistent with the relatively poor results obtained in very recent works when considering a multiple session scenario, e.g. (Arnau-González et al. 2018).

To compare the kind of variability that different sessions introduce in this kind of data, Figure 3 (b) displays the



(a) (b)

Figure 3: Dimensionality reduction by t-SNE on (a) three different sessions for a same subject; and (b) three different subjects for a same session.

data from one of the sessions in (a) along with similar data from two other subjects using the same embedding. As it can be seen, both kinds of variability, across users and across sessions are roughly similar in terms of separability. Inspecting these two plots carefully, we can observe that this separability causes that samples from session 3 of subject 1 (see Figure 3 (a)) appear mixed with samples from session 1 of subject 3 (see Figure 3 (b)). Such an effect may potentially yield incorrect outputs in a biometric identification context when multiple sessions are considered.

CONCLUSIONS

A visual exploratory analysis on several publicly available databases containing annotated EEG signals has been carried out. As a consequence, the extreme variable behavior of this kind of data has been put forward. The arguments provided in this paper are consistent with research results reported in recent literature, and are against the construction of a global model for emotion detection that is valid across the entire population. In addition, results reported in relation to the effect of the session suggest that even a personalized model per user may only be effective along the same session when data was captured, hence imposing a strong limitation on the practical applicability of PSD features from EEG signals for emotion detection. The same principles apply in the user identification case. The separability caused by the session is of a similar magnitude to that caused by the subject, and has a negative effect on the permanence property required in a biometric context.

In view of the results reported in this paper, managing the variability introduced when dealing with different subjects, sessions and equipment seems to be a very hard problem. While most previous research and experiments have been done considering one particular aspect, we suggest that their whole variability must be taken into account in order to significantly improve the performance results obtained to date. In particular, it seems feasible to use data related to the emotion, the session and the identity together, e.g. by designing specific mechanisms to reduce/cancel the subject-related and/or session-related component from the EEG signal, as recently proposed by Arevalillo-Herráez et al. (2019).

ACKNOWLEDGEMENTS

This research has been partly supported by CECOTEC INNOVACIONES, S.L. and the Spanish Ministry of Economy and Competitiveness through projects TIN2014-59641-C2-1-P and PGC2018-096463-B-I00.

REFERENCES

Arevalillo-Herráez M.; Cobos M.; Roger S.; and García-Pineda M., 2019. *Combining Inter-Subject Modeling with a Subject-Based Data Transformation to Improve Affect Recognition from EEG Signals. Sensors*, 19, no. 13. ISSN 1424-8220. doi:10.3390/s19132999.

Arnau-González P.; Arevalillo-Herráez M.; Katsigiannis S.; and Ramzan N., 2018. *On the influence of affect in EEG-based subject identification. IEEE Transactions on Affective Computing*, 1–1. ISSN 1949-3045. doi: 10.1109/TAFFC.2018.2877986.

Arnau-González P.; Arevalillo-Herráez M.; and Ramzan N., 2017. *Fusing highly dimensional energy and connectivity features to identify affective states from EEG signals. Neurocomputing*, 244, 81 – 89. ISSN 0925-2312. doi:http://dx.doi.org/10.1016/j.neucom.2017.03.027.

Ayesh A.; Arevalillo-Herráez M.; and Ferri F., 2014.

Cognitive reasoning and inferences through psychologically based personalised modelling of emotions using associative classifiers. In *IEEE 13th International Conference on Cognitive Informatics Cognitive Computing (ICCI*CC)*. 67–72. doi:10.1109/ICCI-CC.2014.6921443.

Ayesh A.; Arevalillo-Herráez M.; and Ferri F.J., 2016. *Towards Psychologically based Personalised Modelling of Emotions Using Associative Classifiers. International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*, 10, no. 2, 52–64. doi:10.4018/IJCINI.2016040103.

Duan R.N.; Zhu J.Y.; and Lu B.L., 2013. *Differential entropy feature for EEG-based emotion classification.* In *6th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, 81–84.

Hadjidimitriou S.; Charisis V.; and Hadjileontiadis L., 2015. *Towards a Practical Subject-Independent Affective State Recognition Based On Time-Domain EEG Feature Extraction. International Journal of Heritage in the Digital Era*, 4, no. 2, 165–178.

Katsigiannis S. and Ramzan N., 2017. *DREAMER: A Database for Emotion Recognition Through EEG and ECG Signals from Wireless Low-cost Off-the-Shelf Devices. IEEE Journal of Biomedical and Health Informatics.*

Koelstra S.; Muhl C.; Soleymani M.; Lee J.S.; Yazdani A.; Ebrahimi T.; Pun T.; Nijholt A.; and Patras I., 2012. *DEAP: A Database for Emotion Analysis ;Using Physiological Signals. Affective Computing, IEEE Transactions on*, 3, no. 1, 18–31. ISSN 1949-3045. doi:10.1109/T-AFFC.2011.15.

Liu S.; Zhang D.; Xu M.; Qi H.; He F.; Zhao X.; Zhou P.; Zhang L.; and Ming D., 2015. *Randomly dividing homologous samples leads to overinflated accuracies for emotion recognition. International journal of psychophysiology : official journal of the International Organization of Psychophysiology*, 96, no. 1, 2937. ISSN 0167-8760.

Liu Y.; Sourina O.; and Nguyen M.K., 2010. *Real-Time EEG-Based Human Emotion Recognition and Visualization.* In *International Conference on Cyberworlds (CW)*. 262–269. doi:10.1109/CW.2010.37.

Lu Y.; Zheng W.L.; Li B.; and Lu B.L., 2015. *Combining Eye Movements and EEG to Enhance Emotion Recognition.* In Q. Yang and M. Wooldridge (Eds.), *International Joint Conference on Artificial Intelligence (IJCAI)*. AAAI Press. ISBN 978-1-57735-738-4, 1170–1176.

Maaten L.v.d. and Hinton G., 2008. *Visualizing data using t-SNE. Journal of Machine Learning Research*, 9, no. Nov, 2579–2605.

Morris J.D., 1995. *Observations: SAM: the Self-Assessment Manikin; an efficient cross-cultural measurement of emotional response. Journal of Advertising Research*, 35, no. 6, 63–68.

Petrantonakis P. and Hadjileontiadis L., 2010. *Emotion Recognition From EEG Using Higher Order Crossings. IEEE Transactions on Information Technology in Biomedicine*, 14, no. 2, 186–197. ISSN 1089-7771. doi:10.1109/TITB.2009.2034649.

Russell J.A., 1979. *Affective Space is Bipolar. Journal of Personality and Social Psychology*, 37, no. 3, 345.

Salmeron-Majadas S.; Arevalillo-Herráez M.; Santos O.C.; Saneiro M.; Cabestrero R.; Quirós P.; Arnau D.; and Boticario J.G., 2015. *Artificial Intelligence in Education: 17th International Conference, AIED 2015, Madrid, Spain, June 22-26, 2015. Proceedings*, Springer International Publishing, Cham, chap. Filtering of Spontaneous and Low Intensity Emotions in Educational Contexts. ISBN 978-3-319-19773-9, 429–438.

Sohaib A.T.; Qureshi S.; Hagelbäck J.; Hilborn O.; and Jerčić P., 2013. *Foundations of Augmented Cognition: 7th International Conference, AC 2013, Held as Part of HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013. Proceedings*, Springer Berlin Heidelberg, Berlin, Heidelberg, chap. Evaluating Classifiers for Emotion Recognition Using EEG. ISBN 978-3-642-39454-6, 492–501.

Soleymani M.; Asghari-Esfeden S.; Pantic M.; and Fu Y., 2014. *Continuous emotion detection using EEG signals and facial expressions.* In *IEEE International Conference on Multimedia and Expo (ICME), 2014*. 1–6. doi:10.1109/ICME.2014.6890301.

Soleymani M.; Lichtenauer J.; Pun T.; and Pantic M., 2012a. *A Multimodal Database for Affect Recognition and Implicit Tagging. IEEE Transactions on Affective Computing*, 3, no. 1, 42–55. ISSN 1949-3045. doi:10.1109/T-AFFC.2011.25.

Soleymani M.; Lichtenauer J.; Pun T.; and Pantic M., 2012b. *A multimodal database for affect recognition and implicit tagging. Affective Computing, IEEE Transactions on*, 3, no. 1, 42–55.

Wang X.W.; Nie D.; and Lu B.L., 2014. *Emotional state classification from EEG data using machine learning approach. Neurocomputing*, 129, 94–106.

Zheng W.L. and Lu B.L., 2015. *Investigating Critical Frequency Bands and Channels for EEG-based Emotion Recognition with Deep Neural Networks. IEEE Transactions on Autonomous Mental Development*, 7, no. 3, 162–175. doi:10.1109/TAMD.2015.2431497.

MATHEMATICAL MODELLING OF THE RESPONSE OF MALE AND FEMALE SUBJECTS TO VIBRATION UNDER WHOLE-BODY VIBRATION TRAINING CONDITIONS

Naser Nawayseh
Department of Mechanical and Nuclear Engineering
College of Engineering
University of Sharjah
Sharjah, PO Box 27272, United Arab Emirates
nnawayseh@sharjah.ac.ae

Sadeque Hamdan
Sustainable Engineering Asset Management (SEAM)
Research Group
University of Sharjah
Sharjah, PO Box 27272, United Arab Emirates
shamdan@sharjah.ac.ae

KEYWORDS
Model, whole-body vibration training, gender effect.

ABSTRACT

In this work, a model to predict the normalised apparent mass of the standing human body under whole-body vibration training conditions has been developed. The model was a single degree-of-freedom model with a single mass, damping coefficient and stiffness. The parameters of the model were calibrated for a total of 40 subjects (20 females and 20 males) using apparent mass measured over the frequency range 20-45 Hz. For both genders, good agreement was found between the predicted and measured normalised apparent mass. Analysis of the parameters of the model showed gender effect on the damping coefficient but not the stiffness of the model. This indicates that the lower apparent mass of females than males is caused by the higher damping in females than males.

INTRODUCTION

Whole-body vibration (WBV) training machines are widely used in both the medical and sports fields. The practitioner or trainer design a vibration training programme that includes, vibration frequency, magnitude and duration as well as the body posture. The choice of those parameters depends on the medical problem and the muscles targeted in the training (Lienhard et al., 2017). Despite the above, there is contradiction in the scientific literature regarding the effectiveness of using WBV training machines. Some studies reported positive effect while others reported either negative or no effect (Sonza et al., 2015; Yang et al., 2015). This could be partly due to the absence of standards or protocols for a procedure for the experimentation on of WBV training machines. The different procedures and conditions used in the different studies would affect the results (Nawayseh and Hamdan, 2019).

The mechanism by which vibration training benefits the body is not understood. To be able to understand those mechanisms, we need to understand how vibration is transmitted to and through the human body which is referred to as biodynamics responses to vibration. Biodynamic responses to vibration represent how vibration is transmitted to and through the human body (Nawayseh et al., 2019). They are usually represented by the apparent mass (AM) and transmissibility transfer functions. AM and transmissibility have been used to study the effect of different factors (e.g. posture) on the response of the human body to vibration through experimental (Friesenbichler et al., 2014) or modelling studies (Matsumoto and Griffin, 1998). Biodynamic responses have been used to explain the effect of vibration on comfort, health and interference with activities (Griffin, 1990). Hence, understanding the biodynamic responses might help in understanding how the human body benefits from WBV training characterised as exposure to high-frequency high magnitude vibration. Studies have shown differences in body structure and mechanical properties between males and females (Holcombe et al., 2016; Shi et al., 2014). This could vary the transmission of vibration to and through the body and consequently may vary the effectiveness of using WBV training machines between the two genders. For example, a recent study have shown higher AM in males than females but higher floor to head transmissibility in females than males (Nawayseh et al., 2019). This could be one of the reasons for the contradiction in the results among studies as previous studies have used either male or mixed gender subjects.

The objective of this study is to present a single degree-of-freedom mathematical model that represents the AM of the male and female standing subjects under WBV training conditions. The data reported by (Nawayseh et al., 2019) will be used for optimising the parameters.

METHODS

Experiment

The experiment used to calibrate the parameters of the model is described in detail in (Nawayseh et al., 2019). The objective of the experiment was to investigate the effect of gender on the AM during WBV training. Twenty male subjects with average mass 82.8 and twenty female subjects with average mass of 62.7 kg took part in the experiment. A WBV training machine that produces high magnitude of vertical sinusoidal motions was used to produce the vibration. Six frequencies (20, 25, 30, 35, 40, 45 Hz) were used in the experiment. The AM was calculated as the ratio of the root mean squared (r.m.s.) value of the measured force (F) after correction to remove the effect of the mass of the top plate of the force platform to the r.m.s value of the measured acceleration (a) of the vibrating plate (Equation (1)).

$$AM = \frac{\text{r.m.s. } (F)}{\text{r.m.s. } (a)} \qquad (1)$$

Model Description

A single degree-of-freedom model that consists of a mass (m), a damping coefficient (c) and a spring stiffness (k) was developed in order to predict the vertical AM of the standing subjects (Figure 1). The equation of motion of the model with base-excitation was derived using Newton's

second law of motion (Equation (2)). The interface force, F(t) at the base of the model can be calculated using Equation (3).

$$m\ddot{y} + k(y - y_0) + c(\dot{y} - \dot{y_0}) = 0 \qquad (2)$$

$$F(t) = m\ddot{y} \qquad (3)$$

Where y_0 and y are the motion of the base and m, respectively.

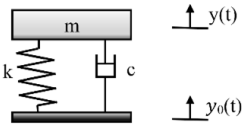


Figure 1: The Proposed One Degree-of-Freedom Lumped Parameter Model

By taking the Laplace Transform of Equations (2) and (3) and substituting ωi for s, the predicted AM in the frequency domain, $M_c(\omega i)$ can be obtained using Equation (4).

$$M_c(\omega i) = \frac{m(c\omega \mathrm{i} + k)}{(m(\omega \mathrm{i})^2 + c\omega \mathrm{i} + k)} \qquad (4)$$

where, ω is the angular frequency in rad/s, and i is $\sqrt{-1}$. The parameters of the model (m, c, and k) were optimised by minimising the summation of the squared error between the measured and calculated normalized AM (NAM) (AM divided by the static mass of the subject) for each subject (Equation (5)). The NAM was chosen instead of the AM in order to minimise the effect of the different masses of the subjects on the parameters.

$$SE = \sum_{i=1}^{7} \left(NAM_m(\omega) - NAM_c(\omega)\right)^2 \qquad (5)$$

Where SE is the squared error over the whole frequency range, $NAM_m(\omega)$ is the measured NAM magnitude as reported in (Nawayseh et al., 2019), $NAM_c(\omega)$ is the predicted AM magnitude. It is well known that the human body behaves like a rigid system at low frequencies (around and below 1 Hz) (Griffin, 1990). So, in order to obtain a reasonable optimised mass that matches the mass of the subjects, the NAM at 0.01 Hz was included in the optimisation process. The measured NAM magnitude at 0.01 Hz was taken as 1.0. The optimization problem was solved using the fmincon function in MATLAB R2018a. The model was used to predict the individual AMs of the 40 subjects as well as their median AM.

RESULTS AND DISCUSSION

With both genders, the single degree-of-freedom model provided a good prediction for the measured median NAM (Figure 2) and the measured individual NAM of the 40 subjects (Figures 3 and 4 for female and male subjects, respectively). A previous study also used a single degree-of-freedom model to predict the response of seated subjects to low frequency low magnitude vibration and reported good prediction for the seated body AM (Toward and Griffin, 2010). This means that whether the body is standing or seated and whether the vibration is high or low in magnitude and frequency, a single degree-of-freedom model is sufficient for predicting the AM. For each subject and with both genders, the NAM predicted by the model at 0.01 Hz was very close to the NAM of 1.0 measured at 0.01 Hz. This supports the observation that the AM of the human body at around 1 Hz is similar to its static mass (Nawayseh and Griffin, 2012).

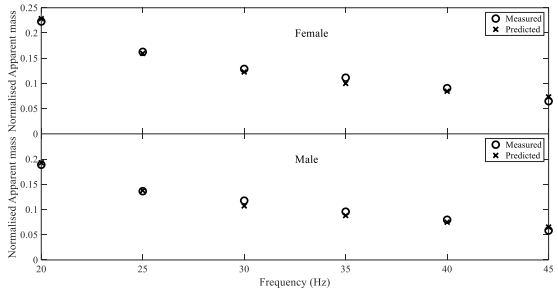


Figure 2: Measured and Predicted Normalised Median AM

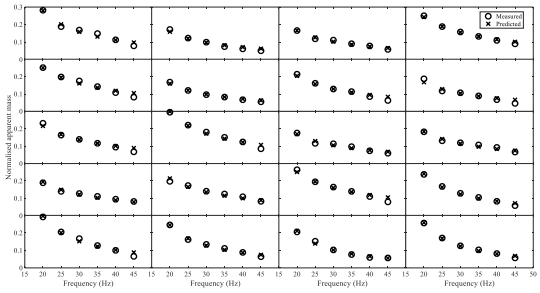


Figure 3: Measured and Predicted NAM of 20 Female Subjects

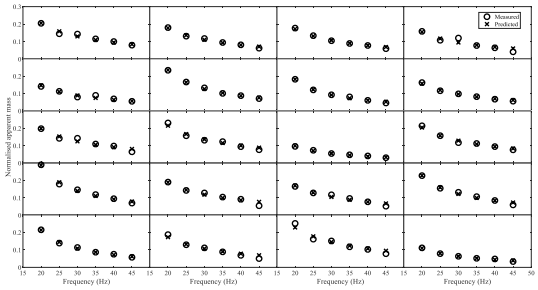


Figure 4: Measured and Predicted NAM of 20 Male Subjects

The parameters obtained for the subjects of both genders are shown in Table 1. Statistical analysis using Shapiro-Wilk test showed that the damping coefficient, c is normally distributed while the stiffness, k is not normally distributed. Hence, the effect of gender on the damping coefficient and stiffness was investigated statistically using parametric and non-parametric tests, respectively. Statistical analysis showed greater damping coefficient in females than males (t-test, $p < 0.05$). However, no statistically significant difference was found in the stiffness between males and females (Mann-Whitney U test, $p > 0.05$). The higher model damping coefficient in females than males is consistent with the female having more body fat than males as reported in (Blaak, 2001). Due to the muscular nature of the body of males compared to the body of females, it was expected that male subjects would have higher model stiffness than female subjects. However, no statistically significant difference was found in the model stiffness between males and females. It has been shown before that the body becomes softer with increasing the vibration magnitude but the softening effect decrease with further increase in vibration magnitude (Nawayseh and Griffin, 2003). It is possible that at the high vibration magnitudes used in the current study, the body reached the limit of softening such that both males and females showed similar stiffness values. It is also possible that female subjects were more anxious when standing on a plate vibrating at high magnitudes and frequencies that they tensed their muscles as a reaction. The results reported in (Nawayseh et al., 2019) showed higher AM for male than female subjects. Higher AM results from higher

static mass and/or higher stiffness and/or lower damping. The mass effect was eliminated by optimising the model parameters using the NAM. The analysis of the damping coefficient and stiffness of the model shown above indicates that female subjects had lower AM than male subjects due to the higher damping in females than males.

Table 1: Optimized Model Parameters for the NAM of 40 female and male Subjects

	Female		Male	
	c [Ns/m]/kg	k [N/m]/kg	c [Ns/m]/kg	k [N/m]/kg
Subject 1	25.85	2179.16	23.38	999.65
Subject 2	17.13	999.47	19.68	999.43
Subject 3	18.08	999.38	18.98	999.52
Subject 4	28.31	1000.17	16.65	999.36
Subject 5	29.32	1000.28	15.48	999.68
Subject 6	17.27	999.47	19.47	2155.69
Subject 7	23.21	999.67	12.66	1969.87
Subject 8	18.34	999.50	16.94	999.35
Subject 9	24.77	999.98	22.45	999.57
Subject 10	28.65	2176.79	24.63	999.97
Subject 11	18.61	999.50	8.63	998.95
Subject 12	20.50	999.46	23.21	999.65
Subject 13	21.88	999.52	18.82	2924.83
Subject 14	24.56	999.86	21.20	999.50
Subject 15	29.41	1000.29	18.64	999.40
Subject 16	17.88	2384.64	18.74	2117.91
Subject 17	22.09	2792.25	15.12	2223.43
Subject 18	18.94	2344.04	19.02	999.61
Subject 19	13.03	2344.00	26.36	999.93
Subject 20	15.92	2754.38	10.25	999.10

CONCLUSION

A single degree-of-freedom mathematical model was developed to represent the NAM of the standing human body under WBV training conditions. The parameters of the model were optimised using NAM of both female and male subjects. The NAM predicted by the model was in good agreement with that measured experimentally for 40 subjects (20 females and 20 males). Analysis of the parameters of the model showed gender effect on the damping coefficient but not the stiffness of the model. This indicates that the lower AM of females than males is caused by the higher damping in females than males.

REFERENCES

Blaak, E. (2001), "Gender differences in fat metabolism.", *Current Opinion in Clinical Nutrition and Metabolic Care*, Vol. 4 No. 6, pp. 499–502.

Friesenbichler, B., Lienhard, K., Vienneau, J. and Nigg, B.M. (2014), "Vibration transmission to lower extremity soft tissues during whole-body vibration", *Journal of Biomechanics*, Vol. 47 No. 12, pp. 2858–2862.

Griffin, M.J. (1990), *Handbook of Human Vibration*, Academic Press, London.

Holcombe, S.A., Wang, S.C. and Grotberg, J.B. (2016), "Modeling female and male rib geometry with logarithmic spirals", *Journal of Biomechanics*, Vol. 49 No. 13, pp. 2995–3003.

Lienhard, K., Vienneau, J., Nigg, S., Friesenbichler, B. and Nigg, B.M. (2017), "Older adults show higher increases in lower-limb muscle activity during whole-body vibration exercise", *Journal of Biomechanics*, Vol. 52, pp. 55–60.

Matsumoto, Y. and Griffin, M.J. (1998), "Movement of the upper-body of seated subjects exposed to vertical whole-body vibration at the principal resonance frequency", *Journal of Sound and Vibration*, Vol. 215 No. 4, pp. 743–762.

Nawayseh, N. and Griffin, M.J. (2003), "Non-linear dual-axis biodynamic response to vertical whole-body vibration", *Journal of Sound and Vibration*, Vol. 268, pp. 503–523.

Nawayseh, N. and Griffin, M.J. (2012), "Power absorbed during whole-body fore-and-aft vibration: Effects of sitting posture, backrest, and footrest", *Journal of Sound and Vibration*, Vol. 331 No. 1, pp. 252–262.

Nawayseh, N. and Hamdan, S. (2019), "Apparent mass of the standing human body when using a whole-body vibration training machine: Effect of knee angle and input frequency", *Journal of Biomechanics*, Vol. 82, pp. 291–298.

Nawayseh, N., Al Sinan, H., Alteneiji, S. and Hamdan, S. (2019), "Effect of gender on the biodynamic responses to vibration induced by a whole-body vibration training machine", *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, Vol. 233 No. 3, pp. 383–392.

Shi, X., Cao, L., Reed, M.P., Rupp, J.D., Hoff, C.N. and Hu, J. (2014), "A statistical human rib cage geometry model accounting for variations by age, sex, stature and body mass index", *Journal of Biomechanics*, Vol. 47 No. 10, pp. 2277–2285.

Sonza, A., Robinson, C.C., Achaval, M. and Zaro, M.A. (2015), "Whole body vibration at different exposure frequencies: Infrared thermography and physiological effects", *Scientific World Journal*, Vol. 2015, p. 452657.

Toward, M.G.R. and Griffin, M.J. (2010), "A Variable Parameter Single Degree-of-freedom Model for Predicting the Effects of Sitting Posture and Vibration Magnitude on the Vertical Apparent Mass of the Human Body", *Industrial Health*, Vol. 48, pp. 654–662.

Yang, F., King, G.A., Dillon, L. and Su, X. (2015), "Controlled whole-body vibration training reduces risk of falls among community-dwelling older adults", *Journal of Biomechanics*, Vol. 48 No. 12, pp. 3206–3212.

AUTHOR BIOGRAPHIES

NASER NAWAYSEH is currently an Associate Professor at the Department of Mechanical and Nuclear Engineering at the University of Sharjah, United Arab Emirates. He obtained his PhD in human responses to vibration from the Institute of Sound and Vibration Research (ISVR) at the University of Southampton in the United Kingdom. After his PhD, he worked as a Research Fellow for three years at ISVR where he was involved in several European and International projects. He then moved to the Gulf Region for an academic position. He is a member of the American Society of Mechanical Engineers (ASME) and the European Society of Biomechanics. His research interests are in the areas of biodynamic responses to vibration, postural stability and seating dynamics.

SADEQUE HAMDAN is currently a Research Assistant at the Sustainable Engineering Asset Management (SEAM) research group. He obtained both his M.Sc. in Engineering Management in 2015 and his B.Sc. in Civil Engineering in 2013 from the University of Sharjah, United Arab Emirates. His research interests are in the areas of air traffic management, maritime transportation, supply chain management, optimization, decision-making and biodynamic responses to vibration.

ENGINEERING SIMULATION

CLASSIFICATION OF ELECTRICAL POWER PEAKS OF PARALLEL OPERATING MACHINES

Armin Siegel
Institute of Material Handling and Industrial Engineering
Technische Universität Dresden
01062 Dresden,
Germany
E-mail: armin.siegel@tu-dresden.de

KEYWORDS
Simulation, power peak determination, power analysis, peak shaving, parallel machines, multi-aisle AS/RS, performance analysis

ABSTRACT

In addition to minimizing the consumption of electrical energy, the determination of power peaks on the demand side poses a challenge. This paper presents a method for estimation of electrical power peak superposition of parallel operating machines using a statistical approach. The method is validated in a simulation study of a high-bay warehouse by a discrete-event material flow simulation. Therefore, the simulation environment was extended by functions for the determination of power demand curves. Furthermore, our method allows the determination of the effect on transport performance by limiting the power peak superposition. It thus opens up new options for power peak control strategies.

INTRODUCTION

In the last two decades the consideration of energy demand and energy efficiency has found it's way into the planning, design and control of technical systems and processes. In addition to the monetary aspects of energy saving, increased efforts by industry made to protect the environment and the climate. Contributed is this development due to legal requirements (see Garwood et al. 2018).
Accordingly, since the mid-2000's, energy and environmenttal aspects have been given greater consideration in the simulation of production and logistic systems, as Wenzel et al. (2018) prove in their publication analysis. The simulation, as an established tool for analysis and evaluation, was extended by these new aspects with different approaches and methods.
In this previous research the power demand and analysis of electrical power peaks are of secondary importance. An increased consideration of power demand is motivated by the following aspects.
Renewable energy sources, such as wind and sun, provide a volatile power supply due to their technological characteristics. The asynchronous fluctuations of power supply and demand have to be compensated by peak load shift and load shaving. The use of energy storage devices can compensate for short-term peak loads. Planning of suitable storage systems requires knowledge of the power peaks occurring in the consumer system (cf. Uddin et al. 2018, Chua et al. 2016).

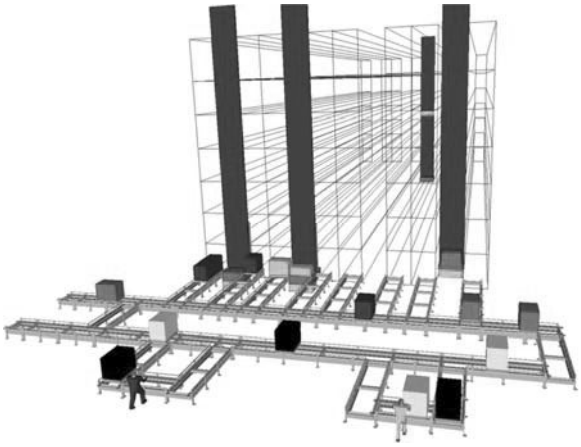


Figure 1: Simulation model of an automated storage and retrieval system (AS/RS) with four parallel operating storage and retrieval machines (SRMs)

A further motivation results from the dimensioning of electrical supply systems (e. g. transformers, power cable cross-sections, circuit breakers), which is based on the maximum loads regularly occurring in the electrical sub-network (see Kiefer 2006). The lack of methods for predicting peak loads therefore leads to an oversizing of the supply systems. This oversizing requires unnecessary resources (Hahn-Woernle and Günthner 2017).
In this paper we focus on the analysis of the electrical power demand for material flow systems. A calculation method for estimating the superposition of power peaks is developed. The validation is later realized with an discrete-event simulation of a high-bay warehouse with an automatic storage and retrieval system (AS/RS) as shown in Figure 1.
The article is organized as follows. First, the emergence and the state of the art for determining power peaks is discussed. Subsequently, the AS/RS used for validation is presented. The power requirement of a storage and retrieval machine (SRM) is presented and the statistical model for determining the superpositions is derived. In the following section, the simulation study for the validation of the model with a discrete-event simulation is presented. The article concludes with a summary and an outlook.

POWER DEMAND OF PARALLEL OPERATING MACHINES

In the planning phase, predicting electrical power peaks of a material flow system is a challenge. Material flow systems have a number of electrical drives that can be operated sim-

ultaneously. Due to the combination of individual components and their function, the measurable maximum power is, in most cases, less than the sum of all rated powers. Figure 2 shows the power supplied from the grid for eight storage and retrieval machines over time. It can be observed that the power curves of the SRMs superimpose and cause a peak power demand of approximately 150 kW, although eight times the maximum power of an SRM would be 280 kW.

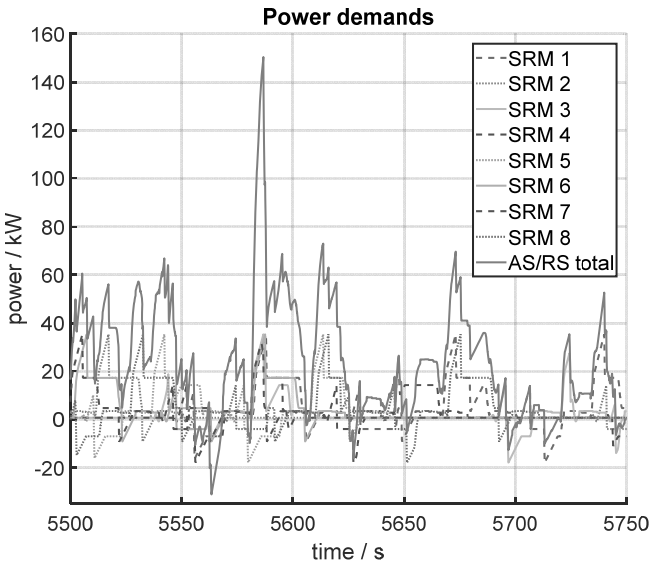


Figure 2: Power demand of an AS/RS with eight SRMs

In the status quo of planning practice, peak power P_d is calculated by the sum of all installed consumers' rated powers P_i multiplied with an estimated demand factor f_d according to DIN VDE 0100-300.

$$P_\mathrm{d} = f_\mathrm{d} \sum_i P_i \tag{1}$$

This takes into account that not all consumers are operated at the same time and describes the amount of power at the connection point (cf. IEEE 1994, Kiefer 2006, Kiank and Fruth 2011). Various table overviews with experience-based approximate statements for certain industries and application areas are available. However, these are usually very coarse and name values of 0.25 for automobile factories or 0.5 for elevator systems, for example (see Kiefer 2006, Kiank and Fruth 2011).

AS/RS system manufacturers report that token release control systems are already being used to reduce maximum power consumption, and respectively, the demand factor. These control systems are not described in the literature. In particular, the effects on system throughput are not known. Hahn-Woernle and Günthner (2017) as well as Hahn-Woernle (2018) are the first using a simulation study to investigate a machine-specific approach for reducing power peaks by delaying the machine start and determining the effects on throughput. The lack of a calculation method to determine the power peaks and the effects on the throughput by using start time delay can be identified.

STORAGE AND RETRIVAL SYSTEMS

Warehouses with AS/RS are used to store goods in a shelf structure. Storage and retrieval machines (SRMs) in such systems receive special attention in the energy and power analysis of material flow systems. Due to their design and operation, these devices have a high energy demand per transport as well as a significant power peak during acceleration. SRMs serve storage locations arranged horizontally and vertically next to each other in a storage aisle. The drives of chassis (horizontal movement) and hoist unit (vertical movement) work simultaneously. Complete warehouse storage systems typically consist of several aisles with simultaneously operating SRMs.

The AS/RS used for validation in this article consists of eight storage and retrieval machines, each operating in its own aisle. The shelves on both sides of each SRM consist of 75 bays 1 m wide and 10 tiers (levels) with a height of 2.4 m each. Furthermore, the kinematic parameters are defined in Table 1.

Table 1: Kinematic parameters of the SRMs

	Chassis (c)	Hoist unit (h)
Jerk r / m · s^{-3}	2.0	1.2
Acceleration a / m · s^{-2}	0.5	1.0
Velocity v / m · s^{-1}	3.0	1.0

There are two different operating modes for a SRM, as shown in Figure 3. On one hand there are the single cycle movements (Fig. 3 left), where one load is picked up from the transfer point (I/O) and transported to it's storage location. After completion of the transport operation, the SRM returns empty (operation 1 and 2). During retrieval, the SRM moves empty to the load, picks it up and returns to the transfer point (operation 3 and 4). In contrast, at double cycles, storage and retrieval operation take place in a joint movement (Figure 3 right).

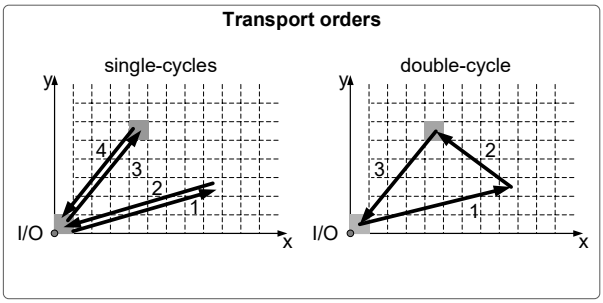


Figure 3: Shelf operation cycles of an storage and retrieval machine (SRM)

The maximum throughput, i. e. the number of pallets per hour that can be stored or retrieved, is defined by FEM 9.851. The regarded AS/RS achieves a maximum throughput of 54.7 pallets with 109.4 movements per hour in single cycle operation for each SRM. When operating in double cycles, the maximum throughput of almost 67 pallets with slightly more than 100 movements per hour could be realized. The calculations include a load handling time of 12 s.

Power demand of SRMs

A power demand model for SRMs was developed by the author in earlier projects and validated by several real system measurements (Siegel et al. 2013; 2018). The physical model is comparable with the approaches in Ertl (2016) and Braun (2016). A kinematic motion model for the chassis and the hoist unit has been developed. With limitation of the maximum deviation of the acceleration (jerk), a seven-phase motion model has been constructed. The graphs of the motion variables acceleration, velocity and distance result respectively from integration of jerk. Figure 4 above shows the trajectory of the kinematic values for a horizontal movement of one validation SRM.

From the trajectory, the required mechanical power is determined by consideration of moved mass and drive resistance. Mechanical losses, e. g. the driving resistance of the wheel-rail combination, and electrical losses, e. g. in the frequency inverter as well as of the powertrains are included in the model calculations. The power requirement for a complete movement results from the superposition of the horizontal and vertical motion. In addition, a basic electrical power loss is added. Figure 4 below shows the power demand for chassis P_c and hoist unit P_h as well as the resulting total demand P. For the validation system a maximum power peak of approximately 35 kW results.

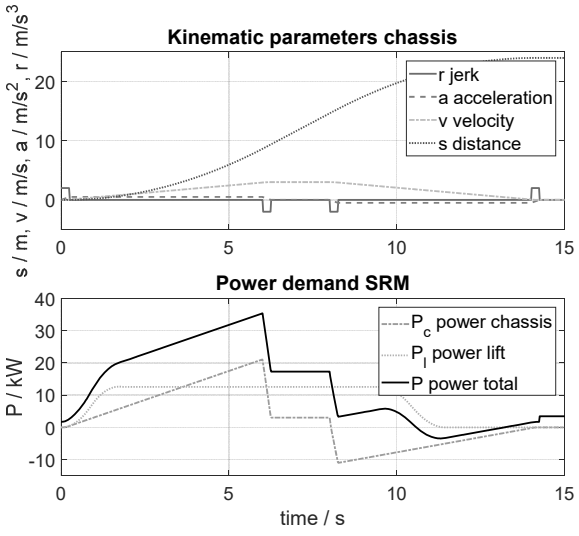


Figure 4: Motion trajectory for chassis and power demand for chassis and hoisting unit of a SRM

Statistical power peak superposition model

The probability for superposition of the SRMs acceleration power peaks is estimated with the statistical model. The approach is based on the assumption that the maximum power requirement is demanded during the entire acceleration time. As Figure 4 below shows from zero to six seconds, this is only the case for a point of the acceleration process. Additionally prerequisite for the occurrence of a maximum power peak are a sufficiently long chassis movement and that the hoist is raised for the duration of the chassis acceleration process. Therefore, the assumption provides an estimate of the upper limit for the occurrence of power peaks.

Due to the fixed parameters for chassis acceleration a_c and velocity v_c (see Table 1), a maximum time requirement t_{acc} for the acceleration process is determined. If the number of movements n_m for a SRM is known, the probability p of being accelerated can be calculated for a given time interval T.

$$p = \frac{n_m \cdot t_{acc}}{T} \quad \text{with} \quad t_{acc} = \frac{v_c}{a_c} \tag{2}$$

The next step is to determine when two or more SRMs are simultaneously in the acceleration phase. In an example of $n = 3$ SRMs there are $2^n = 8$ possible permutations, where none, one or more machines are in the acceleration phase. Under the assumption that

- the probability p of being accelerated will be the same for all SRM and
- accelerating of one SRM is independent of others,

the probability that in an AS/RS with n SRMs k will start simultaneously can be determined. Therefore, the binomial coefficient according to Equation (3) is applied (see Thies 2013).

$$p_k = \binom{n}{k} p^k (1-p)^{n-k} \tag{3}$$

The probability p describes that a machine accelerates (see Equation (2)). (1-p) is the counter-probability and determines the probability that the machine does not accelerate.

AS/RS simulation model

A simulation model is used for validation. It depicts the described pallet high-bay warehouse and was created with the AutoMod simulation environment. Storage and retrieval of pallets into the warehouse take place independently of each other. An average number of pallets per hour is specified as the target value per SRM. The destination storage compartments are chosen randomly from the compartments of an aisle. A compartment allocation is not considered.

The discrete-event simulation environment had to be extended by continuous calculation methods for the determination of the power demand. Therefore, the equations of the energy-movement model were implemented in Matlab and then compiled into a DLL (Dynamic Link Library). This calculation DLL was integrated into the simulation model of the material flow system using the C interface of AutoMod (see Siegel et al. 2018)

EXPERIMENTS

The simulation experiments are presented below. First, single cycles are examined and evaluated. The investigation is extended to double cycles combined with single cycles in the second step.

Single cycle operation

Initially, only storage (retrieval) operations are considered in isolation (see Fig. 3 left). Thus, for a given throughput, the

number of movements is defined as twice this. For the example system presented, a throughput of 30 pallets per hour and one SRM is assumed. This results in 60 movements per hour for one SRM. The acceleration time t_{acc} is calculated with 6 s and it follows a probability for an acceleration state at time unit $T = 3600$ s of 10 % (see Equation (2) and Table 1). The results of Equation (3) are summarized in Table 2 and shows that 43 % of time the machines are not in the start-up process.

Table 2: Theoretic probability of parallel start-ups

Nr. k of (parallel) machine starts	probability
0	0.4305
1	0.3826
2	0.1488
3	0.0331
4	0.0046
5	0.0004
6	$2.2 \cdot 10^{-5}$
7	$7.2 \cdot 10^{-7}$
8	$0.1 \cdot 10^{-7}$

In addition, it can be seen that for the simultaneous start-up of four or more SRMs, a probability of 0.46 % or less is determined. Together, there is a probability of approximately 0.5 % for parallel starting of more than three SRMs. This means that 18 s per hour a power peak of more than three times the maximum power consumption of one SRM can occur. It is worth noting that the probabilities of six or more SRMs accelerating at the same time is close to zero.

With the validation system, a simulation of 10 hours of warehouse operation was carried out. First, the frequencies of the acceleration superpositions based on the start time alone of an SRM are compared with the statistical model. In a second evaluation, only those movements with a maximum power peak of 35 kW are taken into account.

Figure 5 shows the probabilities and relative frequencies that a simultaneous start process of zero to eight SRMs has taken place for the approaches. It becomes apparent that the probabilities of the statistical model correspond very well with the start based relative frequencies from simulation. If the actual movements with maximum power peak (35 kW) are taken as a base, the relative frequencies for $k = 1 \ldots 8$ are less. This can be explained by the fact that a maximum power peak occurs for single cycles at the highest in 50 % of all movements.

For single cycles it only happens during the movement into the shelf that a chassis movement and hoisting takes place at the same time. The percentage further reduced by the fact that among these 50 % are movements that run only without hoisting at the level of the transfer point (in the example approximately 10 % of all movements). In addition, there are movements that have too short driving distance or too low lifting height to cause a maximum acceleration power peak (in the example at least about 18 m distance and 6 m hoisting height). Of the 60 movements per hour in the simulation, only 16 have a peak of 35 kW which corresponds to a share of almost 27 %.

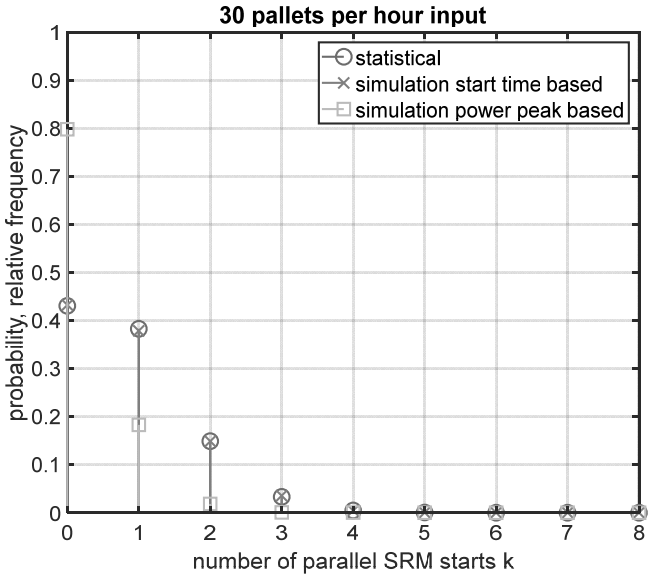


Figure 5: Comparison statistic model and simulation for single cycle operation and 30 pallets throughput per SRM

In the simulation records the power peak of 150 kW, shown in Figure 2 results. The power demand from parallel acceleration of five SRMs superimposes. The machines one, three, four and six are involved with a maximum peak acceleration and SRM seven with an acceleration peak of about 31.5 kW. The number of pallets was then increased in steps of 5 to the theoretical maximum of 55 pallets (110 movements) per hour. Figure 6 shows the results.

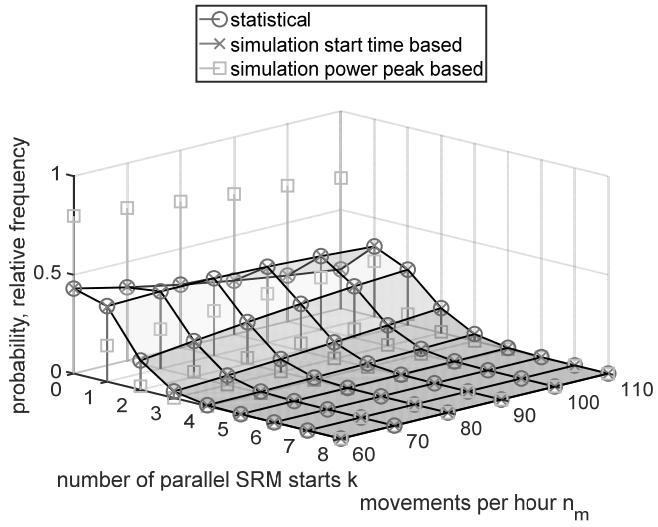


Figure 6: Comparison statistic model and simulation for single cycle operation of SRMs

The probabilities and relative frequencies for simultaneous starting over $k = 0 \ldots 8$ and the number of movements are plotted. All experiments show a good correlation between the statistical model and the start time based trip evaluation of the simulation model. In addition, the behavior is basically similar to that of the example discussed above. There is an increase in the probabilities and relative frequencies for simultaneous starting of two to eight machines with the number

of movements. Although, the probabilities and relative frequencies for more than four devices increase only insignificantly. The probabilities and relative frequencies that none of the machines will accelerate decreases as the number of trips increases. In relation to this, the movements in which a maximum power peak was registered behave similarly, as in the case of 60 movements.
The power peaks occurring in the simulation are summarized in Table 3. In it, the number of power peaks indicates a correlation with the number of trips, whereas the height of the peaks does not show a distinct correlation. For example, in the case with the maximum number of movements, only a maximum peak of 143 kW can be determined. In contrast with 60 movements per hour, the second highest power peak of 150 kW is generated. Altogether it can be stated that the values recorded with the simulation are slightly more than half of the theoretically maximum possible power peak of 280 kW (8 · 35kW).

Table 3: Number and maximum of power peaks depending on movements per hour of 10 hours warehouse simulation for single cycle operations

Movements per hour	Power peaks > 120 kW	Maximum in kW
60	1	150
70	4	137
80	8	138
90	13	150
100	31	153
110	35	143

Combined double and single cycle operation

In double cycles the evaluation of movements from the throughput is much more challenging, since a work cycle consists of three movements. The second movement depends on the storage compartment locations and cycle direction on the shelf (see Figure 3). Whereby it can be assumed by means for the second movement that with approximately the same number of movements or 50 % probability, driving and lifting as well as driving and lowering is done. The proportion of movements with driving and hoisting includes these for which the same restrictions (e. g. short distance) as for single cycles apply. In addition, there are movements between storage locations where only a lifting movement is performed.
Storage operations that consist of single and double cycles thus are most challenging in determining the movements with maximum power peak in advance. Depending on the throughput, this means one and a half to two times the number of movements. Compared to an operating mode of only single cycles, the operation with single and double cycles reduces the movements for a given throughput by up to a quarter.
The statistical model assumes that each trip can cause a maximum power peak and derives the number of movements from the assumption of single cycles. The model can thus be used to estimate the maximum power peaks that occur. In particular, it is possible, because the movements with maximum power peaks are less in the real system.

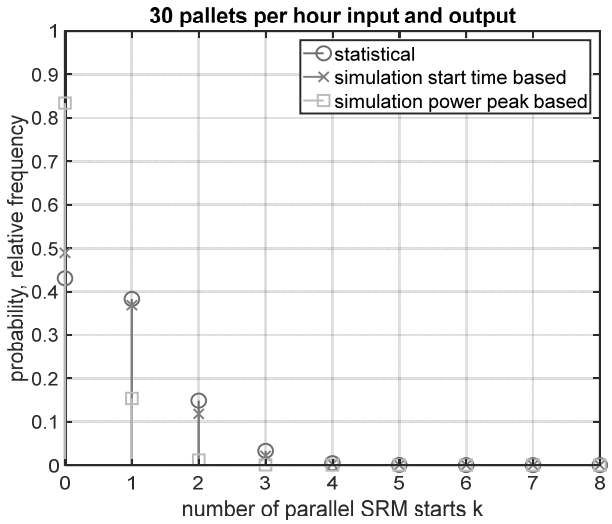


Figure 7: Comparison statistic model and simulation for combined single and double cycle operations as well as 30 pallets throughput per SRM

An experiment with a throughput of 30 pallets per hour for storage and retrieval was carried out. Figure 7 shows the results. A total mean of 51 movements per hour is carried out by one SRM in the simulation model. The 51 trips per hour include about 9 double cycles, which corresponds to a share of 15 % and explains the difference with the expected 60 movements in the case of single cycles. On the other hand, only double cycles would result in 45 movements per hour.
This behavior is the reason why the statistical and simulation model based on start times no longer matches as well as in the case of single cycles. Fewer starts are performed in the simulation model as calculated in the statistical model, which results in a lower relative frequency of simultaneous start processes. The corresponding relative frequency that the machine dose not accelerate is higher. From the movements performed, those with maximum power peak start-ups occur less frequently, which is why the peak based relative frequencies for $k = 1 \ldots 8$ are lower than the probability of the statistical model and the relative frequency of the start time based counting. The opposite happens in case of $k = 0$.
The investigations were also carried out with different throughput in the case of combined cycles. Therefore, the throughput was increased in steps of 5 pallets per hour. For the assumption of single cycles, this means an increase in movements of 10 and is comparable to the single cycle analysis. Figure 8 shows the results of the comparison.
The probabilities and relative frequencies are plotted equivalently as in the case of single cycles. The movements are noted for the assumption of single cycles. This information was supplemented by the actual movements executed in the simulation model in brackets.
A basically similar behavior to the discussed case with throughput of 30 pallets per hour could be observed. In addition, it can be remarked that with increasing throughput the share of double cycles increases up to a set throughput of 70 pallets per hour (140 single cycles movements), were only double cycles are executed. In this case, 101 movements are realized and 67 pallets are transported, which is approximately the theoretical maximum according to FEM 9.851.

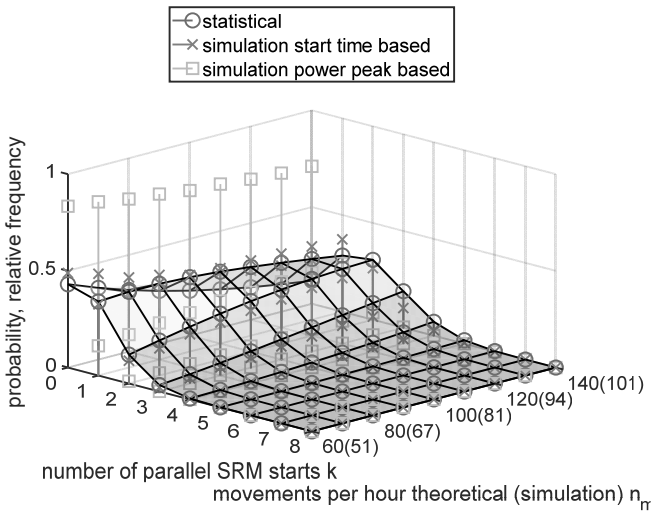


Figure 8: Comparison statistic model and simulation for combined single and double cycle operations of SRMs

In addition, as for single cycles, with an increase in throughput, there is an increase in the probabilities and relative frequencies that SRMs are simultaneously in the start-up process. For example, with 70 pallets (140 single cycle trips), the cumulative probability that more than three SRMs will start simultaneously increases to slightly more than 9 % or a little less than five and a half minutes per hour. Similarly, with increasing movements there is a decrease in the probability that none of the machines will start-up. The probability that one machine accelerates has a maximum between 90(73) and 100(81) movements. The differences between the models can also be explained by the different number of movements.

The power peaks occurring in simulation were also noted for combined cycles. Table 4 summarizes them. In the case of power peaks bigger than 120 kW, there is also a tendency for these to occur more frequently as the number of trips increases. For the height of the peaks there is no obvious correlation. The highest peak occurs with 160 kW at theoretical single cycles with 110 respectively 87 in simulation recorded movements per hour. For this peak, the start-up processes of four and the constant drive of two SRMs superimpose each other.

Table 4: Number and maximum of power peaks depending on movements in 10 hours warehouse simulation for combined single and double cycle operations

Movements per hour	Power peaks > 120 kW	Maximum in kW
60 (51)	0	113
70 (57)	4	132
80 (67)	6	150
90 (73)	3	148
100 (81)	5	145
110 (87)	8	160
120 (94)	5	133
130 (98)	14	150
140 (101)	12	134

CONCLUSION AND OUTLOOK

In this article, we presented a method for determining the probability of power peaks in systems with parallel operating machines. For these systems, the acceleration processes to reach the maximum speed, which causes the highest power peak of a storage and retrieval machine, were chosen. Using a developed statistical model, the probability of machines being started at the same time can be determined.

The simulative validation of the developed model shows good suitability for determining the probability of the upper limit for simultaneous start-ups. An AS/RS with eight parallel SRMs was used as the validation example. It can be concluded that the start-up superposition of an increasing number of machines becomes less likely however the occurrence of the power peak magnitude shows no obvious correlation.

So far not discussed is handling parallel start-up superposition probabilities in the statistical model near zero. Similar to the common approach with the demand factor, a lower limit for the probability of power magnitude occurrence has to be defined for the practical application and the remaining probability has to be accepted or intercepted by e. g. a control system. In summary, the presented method improves the approach with the demand factor. In particular, with the statistical model it is possible to determine the factor on an individual system basis.

The statistical model can be further improved in different ways. One possibility is to determine the movements with simultaneous driving and hoisting more precisely. Nevertheless, it is necessary to ensure that the determination is an estimation of the upper limit for superposition. Another possible improvement is considering only the time portion where the power demand of the start-up exceeds that of a given value (e. g. the demand of constant drive and hoisting).

In addition, it should be noted that not only the power demand of the acceleration phases contributes to the system power peak. Also constant driving and lifting causes a power requirement. Therefore, the maximum power limit should be carefully selected.

Furthermore, the presented model allows determination of the superposition time fraction of several start-up processes. Thus, the effects for the avoidance of power peak superposition can be calculated and made transparent for the user. This opens the option to prevent the superposition in the system control (material flow control) e. g. by start shifting in order to limit the power demand. Additionally this control approach allows handling special events such as complete system start-up.

The modification of the start times (and the kinematic parameters) of the device trajectories provides an appropriate way of power smoothing (see Windmann et al. 2018, Cardenas et al. 2009). As a result, power peak control can contribute to power peak shaving and load shifting for material flow systems. Subsequently in further research, we will investigate the impact on smoothing power consumption from the grid by integrating a power peak control as well as an energy storage system in our simulation model.

ACKNOWLEDGEMENT

I thank Karsten Turek for the support and for comments that greatly improved the paper. As well I thank David Jones for the proofreading.

REFERENCES

Braun M.S.A., 2016. Entwicklung, Analyse und Evaluation von Modellen zur Ermittlung des Energiebedarfs von Regalbediengeräten. Ph.D. thesis, Karlsruher Institut für Technologie, Karlsruhe.

Cardenas J.J.; Garcia A.; Romeral J.L.; and Andrade F., 2009. A genetic algorithm approach to optimization of power peaks in an automated warehouse. In 2009 35th Annual Conference of IEEE Industrial Electronics. IEEE, Porto, 3297–3302.

Chua K.H.; Lim Y.S.; and Morris S., 2016. Energy storage system for peak shaving. International Journal of Energy Sector Management, 10, no. 1, 3–18. ISSN 1750-6220.

DIN VDE 0100-300, 1996. Errichten von Starkstromanlagen mit Nennspannungen bis 1000 V, Beuth.

Ertl R.F., 2016. Energiebedarfsermittlung und Energieeffizienzbewertung von Regalbediengeräten in automatischen Kleinteilelagern. Ph.D. thesis, TU München, München.

FEM 9.851, 2001. Performance Data of S/R Machines Cycle Times, FEM.

Garwood T.L.; Hughes B.R.; Oates M.R.; O'Connor D.; and Hughes R., 2018. A review of energy simulation tools for the manufacturing sector. Renewable and Sustainable Energy Reviews, 81, 895–911. ISSN 13640321.

Hahn-Woernle P., 2018. Reduzierung der Betriebskosten mehrgassiger Lagersysteme mittels elektrischem Lastmanagement. phdthesis, TU München, München.

Hahn-Woernle P. and Günthner W.A., 2017. Power-load management reduces energy-dependent costs of multi-aisle mini-load automated storage and retrieval systems. International Journal of Production Research, 1, no. 1, 1–17. ISSN 0020-7543, 1366-588X.

IEEE, 1994. IEEE Recommended Practice for Electric Power Distribution for Industrial Plants, IEEE. doi:10.1109/IEEESTD.1994.121642.

Kiank H. and Fruth W., 2011. Planungsleitfaden für Energieverteilungsanlagen: Konzeption, Umsetzung und Betrieb von Industrienetzen. Publicis Publ, Erlangen. ISBN 978-3-89578-359-3. OCLC: 697813857.

Kiefer G., 2006. VDE 0100 und die Praxis: Wegweiser für Anfänger und Profis. VDE-Verl, Berlin, 12. ed.

Siegel A.; Turek K.; Michelini E.; and Schmidt T., 2018. Hybrid modeling approach for prediction of energy demand and power peaks in intralogistic systems. In 24. Symposium Simulationstechnik. ARGESIM Verlag, Hamburg, 56, vol. 168, 81–88.

Siegel A.; Turek K.; Schmidt T.; Schulz R.; and Zadek H., 2013. Modeling the energy need of storage and retrieval vehicles and different storage operating strategies for the reduction of the energy need. Logistics Journal, 2013.

Thies K.D., 2013. Elementare Einführung in die Wahrscheinlichkeitsrechnung, Informationstheorie und stochastische Prozesse mit Warteschlangentheorie für Computernetzwerke: mit einer wahrscheinlichkeitstheoretischen Leistungsanalyse des Ethernet. Berichte aus der Informationstechnik. Shaker, Aachen, 2. ed.

Uddin M.; Romlie M.F.; Abdullah M.F.; Abd Halim S.; Abu Bakar A.H.; and Chia Kwang T., 2018. A review on peak load shaving strategies. Renewable and Sustainable Energy Reviews, 82, 3323–3332. ISSN 13640321.

Wenzel S.; Peter T.; Stoldt J.; Schlegel A.; Uhlig T.; and Josvai J., 2018. Considering energy in the simulation of manufactoring systems. In 2018 Winter Simulation Conference (WSC). IEEE, Gothenburg, Sweden, 3275–3286.

Windmann S.; Niggemann O.; and Stichweh H., 2018. Computation of energy efficient driving speeds in conveying systems. at - Automatisierungstechnik, 66, no. 4, 308–319. ISSN 2196-677X, 0178-2312.

APPLICATION OF MACHINE LEARNING TO MODEL A BIOLOGICAL REACTOR IN A WASTEWATER TREATMENT PLANT

Jan Studziński, Andrzej Ziółkowski
Systems Research Institute, Polish Academy of Sciences, 01-447 Warsaw, Poland
studzins@ibspan.waw.pl, a.ziolkowski1@gmail.com

KEYWORDS
Management of the water supply company, pump control, hydraulic optimization of the water supply network.

ABSTRACT

The paper presents examples of energy and pressure optimization in water supply networks. Optimization tasks are defined as multi-criteria optimization tasks. The logic of fuzzy sets is used to define quality criteria in these tasks. The control variables in the optimization calculations are the rotational speeds of pumps in the water intake pumping stations of the water supply network. Genetic algorithm is used for optimization calculations. Optimization tasks are solved by the MOSKAN-W IT system developed at the System Research Institute of the Polish Academy of Sciences in Warsaw. This system was developed for the water supply company GPW S.A. in Katowice and the given calculation examples concern the water supply network of this company..

INTRODUCTION

The System Research Institute has been developing IT systems supporting the management of urban water supply networks for over a dozen years. The result of this work is the MOSKAN-W system designed for modelling, optimization and design of water networks. The system supports the management of water supply networks in the scope of flow and pressure regulation, selection of pipeline diameters, operation of pumps, detection and location of failures and determination of optimal chlorine doses to improve the quality of produced water. The need to reduce the operating costs of the water network, increase the efficiency of water supply system management and support the investment process are the problems that every water supply company has to face. The MOSKAN-W system was developed to achieve these objectives with the use of modern information technologies employed to develop Internet applications. This solution opens up the possibility of remote operation and expansion of the system with access functions from mobile devices. The MOSKAN-W system is a key module of a wider IT system consisting of a GIS system and SCADA, AMR and billing monitoring systems. The GIS system is integrated with monitoring systems. In turn, the integration of MOSKAN-W with a GIS system simplifies the process of data preparation for modelling and optimization of the water supply network under consideration. The integration of both systems makes it possible to export from the GIS system to MOSKAN-W a hydraulic graph of the network or a specific part of it and automatically start the calculations. The MOSKAN-W system was developed for the waterworks enterprise GPW S.A. in Katowice. The water supply network of this enterprise covers the whole area of the Upper Silesian Voivodeship. It is a distribution network supplying drinking water to the cities of Upper Silesia and their water supply systems.

OPTIMISATION TASKS PERFORMED BY THE MOSKAN-W SYSTEM

The innovation of the developed IT system is supporting the management of a water supply network by means of optimization algorithms. In order to define the optimization task, it is necessary to define control (decision) variables and the scope of their variability as well as criteria for assessing the quality of the solution found (quality criteria=target functions). Optimization tasks occurring in the management of water supply networks are in most cases multi-criteria optimization tasks. A convenient mathematical apparatus for describing such tasks are fuzzy sets and fuzzy logic. All criteria and limitations can then be treated as fuzzy values forming a fuzzy set. A certain range of values of a defined quality criterion may meet fully accepted requirements (preferred values), but a wider range of criterion values may also be given in which these requirements are met to a lesser extent (permissible values). There is also a certain range of criterion values where the requirements are not fulfilled at all. The use of fuzzy sets facilitates the unification of the problem description, the implementation of the program interface and the development of optimization algorithms. Criteria and limitations represented by fuzzy sets can be approximated, for example, by a simple trapezoidal function, defined by giving four numbers (Fig. 1).

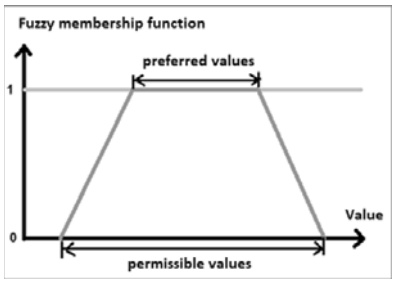


Fig. 1. Trapezoidal affiliation function defined in fuzzy sets.

The trapezoidal function in Fig. 1 defines the degree of belonging to the set 'Criterion fulfilled' of the values of the defined quality criterion. For preferred values the degree of belonging is 1, and for values outside the range of permissible values it is 0. For other values of the criterion the function takes values from the range (0.1). To determine

the optimal solution, it is necessary to find control variable values for which the degree of belonging of the values of all quality criteria defined in the optimization task to the value 'Criterion fulfilled' will be 1, and if this is not achievable, then control variable values for which the degree of belonging of the value of the worst criterion will be the highest. The MOSKAN-W system shows the extent to which each criterion belongs to a set of preferred or limit values and the values of the control variables of the best solution found after the optimization calculations have been completed. Gradient optimization methods fail in difficult optimization tasks, which include water supply network management tasks. Such tasks can be solved with the use of genetic optimization methods. Algorithms of genetic optimization guarantee approach to optimal solutions with a sufficient number of calculations of functions defining quality criteria, although they do not guarantee finding an optimal solution. The principle of genetic algorithms can be explained by analogy to the phenomena of population adaptation to the environment in biological systems. The population is made up of individuals whose characteristics are determined by a set of genes (genotype). Depending on their characteristics, the individuals are better or worse adapted to the environment. During the evolutionary adaptation of the population to the environment, the following phenomena occur: some of the worst adapted individuals are removed from the population (the extinction of the weakest); the population is supplemented by new individuals whose genotypes are created by mixing parental genotypes or changing gene parts (mutations).

A genotype is a set of all control variables describing a water supply network (e.g. diameter, pressure, flow), a gene is one parameter from this set (e.g. pressure), and an individual is each set of values of all control variables. The calculated value of the quality criterion determines the degree of adaptation of a given individual from the generated population of individuals. The process of evolutionary adaptation of the population to the environment allows to find solutions that best meet the adopted quality criteria (the best adapted individuals). Optimization calculations are carried out in two stages. In the first stage, different individuals are randomly generated, which are added to the population until a given population size is reached. At this stage, individuals are not removed from the population (no extinction of the weakest). In the second stage, in each iteration new individuals are added and the worst adapted individuals are removed. Thus, the size of the population is constant and its total adaptation cannot decrease. Additionally, each gene can be mutated with some probability. This means that in the genotype of a new individual, part of the genes may not come from the genotypes of individuals in the population, but may be generated randomly.

PUMP CONTROL ALGORITHM TO REDUCE ENERGY COSTS LOCALLY

The task of reducing the costs of energy used by pumps in water intake pumping stations of the water supply network does not take into account the costs of water acquisition and treatment. The optimisation task is defined by taking as the main quality criterion the minimisation of energy costs consumed by pumps in the whole network or in a selected part of it. The control variables are the pump speeds. In this task, additional criteria defining the required pressure values in selected network nodes can be applied. Considering the pressure criterion in the task of pump control means hydraulic optimization of the water supply network. Below we will show an example of a solution to the task of energy-saving control of pumps in a selected fragment of the water supply network of GPW S.A., so the calculations are local in nature. Fig. 2 shows the screen of the MOSKAN-W program used to start the pump control algorithm. The data entered at this stage of the calculation concern: a preset calculation option (energy, i.e. minimizing energy consumption), a preset electricity price (0,23 PLN/kWh), a simulation time for network operation (24 hours), and the network load multiplier (1, i.e. the nominal network load is taken into account).

Fig. 2. Starting the pump control algorithm.

The next stage of calculations is defining the basic criterion in the task of optimization, i.e. energy criterion, using the logic of fuzzy sets. The criterion is defined by entering 4 values of the parameter 'energy': Min., Od, Do, Max (Fig. 3). In the given example, the given function of belonging the value of this criterion to the fuzzy set is a triangular one, i.e. Min=Od=Do=0 and Max=4.000 PLN (Fig. 6). It means that the value of the criterion has the value of 1 with zero energy costs and is decreasing linearly to 0 with energy costs equal to 4.000 PLN, i.e. in the defined criterion there are no values of preferred energy costs but only acceptable values.

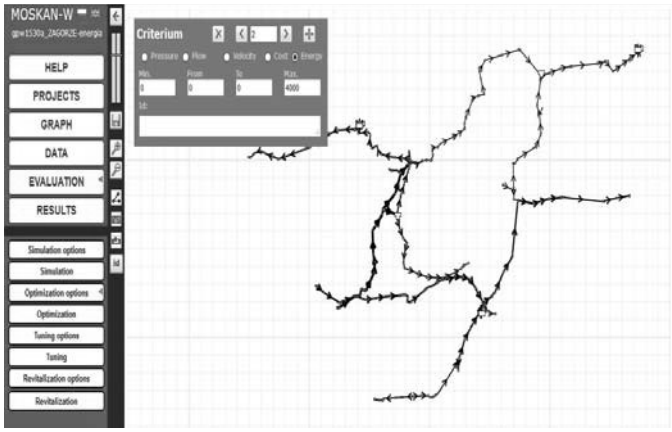


Fig. 3. Definition of the energy quality criterion.

As mentioned above, additional pressure quality criteria can be introduced in the task of reducing energy costs (Fig. 4). In the calculated example, the required pressure range for three selected nodes of the water supply network was defined: 11558, 13912, 7225 (Fig. 4). In this case the preset function of belonging the value of the pressure criterion to the fuzzy set is trapezoidal, i.e. Min=1 atm., Od=1,5 atm., Do=5 atm. and Max=10 atm. (Fig. 6). This means that the preferred values of pressure in these nodes are within the range from 1,5 to 5 atm., while the acceptable values of pressure are within the range of 1-1,5 atm. and 5-10 atm.

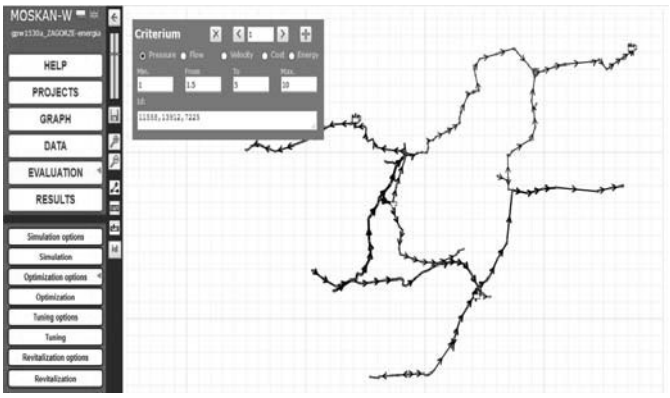


Fig. 4. Defining a pressure quality criterion for selected network nodes.

The control variables (decision variables) in the optimization task are the pump speeds. Fig. 5 shows the definition of acceptable speed variation ranges for four out of five pumps operating in a selected part of the water supply network and included in the algorithm as control variables (pumps 21142, 21202, 22035, 21140, 22048). The range of permissible values for each control variable can be defined individually, e.g. for pump 21142 it is 1 – 1,2, for pump 21140 it is 0,9 – 1, while value 1 is the nominal speed of the pump according to its pressure characteristics.

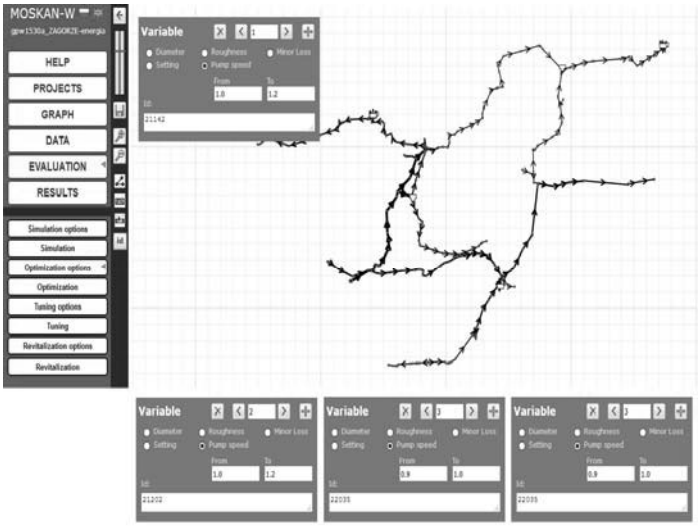


Fig. 5. Definition of speed variation ranges for four pumps.

The optimization results obtained for the two quality criteria, energy and pressure, and the five control variables are shown in Fig. 6, while the optimization results for only one energy criterion and also the five control variables are shown in Fig. 7.

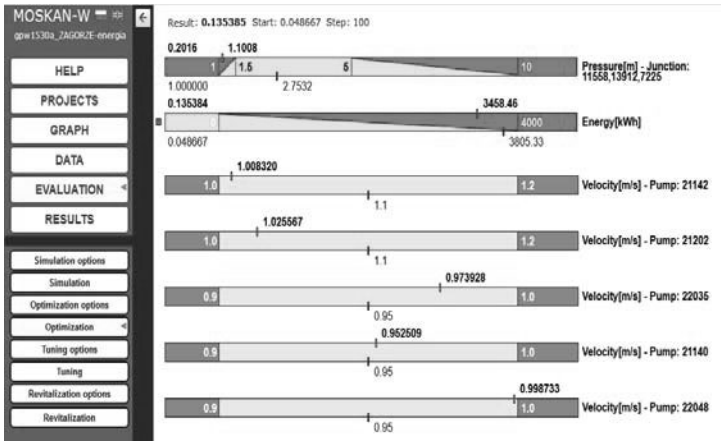


Fig. 6. Calculation results for two quality criteria and five control variables.

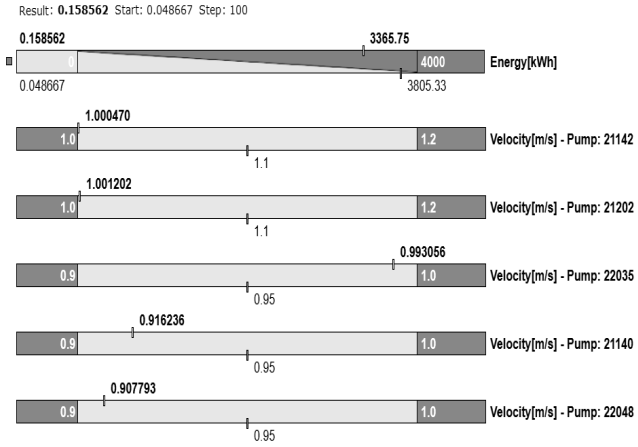


Fig. 7. Calculation results for one quality criterion and five control variables.

The evaluation of the effectiveness of the presented pump control algorithm on a local scale is as follows: In the first calculated example with two quality criteria, the pressure in the three selected network nodes decreased and is now 27% lower than the preferred minimum pressure of 1,5 atm. At the same time, the cost of energy consumed by the five pumps was reduced by 9% as a result of the optimisation. The omission of the pressure criterion in the second example reduced energy costs by more than 11%.

PUMP CONTROL ALGORITHM TO REDUCE OPERATING COSTS GLOBALLY

The network model enables the study of different variants of network operation and the assessment of the impact of the choice of a specific variant of operation on the operation costs of the water supply network. The operating costs are understood as the costs of electricity used by the pumps in the water intake pumping stations of the network and the costs of water treatment in these stations. In particular, the impact on the operation of the network of shutting down selected pumping stations or increasing or decreasing the amount of water pumped by them may be examined. The results of such analysis may indicate the direction of changes most strongly affecting the reduction of costs. Since many technological limitations and various limitations related to water quality assurance have to be taken into account when

selecting variants, full automation of the optimization process is very difficult and determining simplified solutions is not very useful.

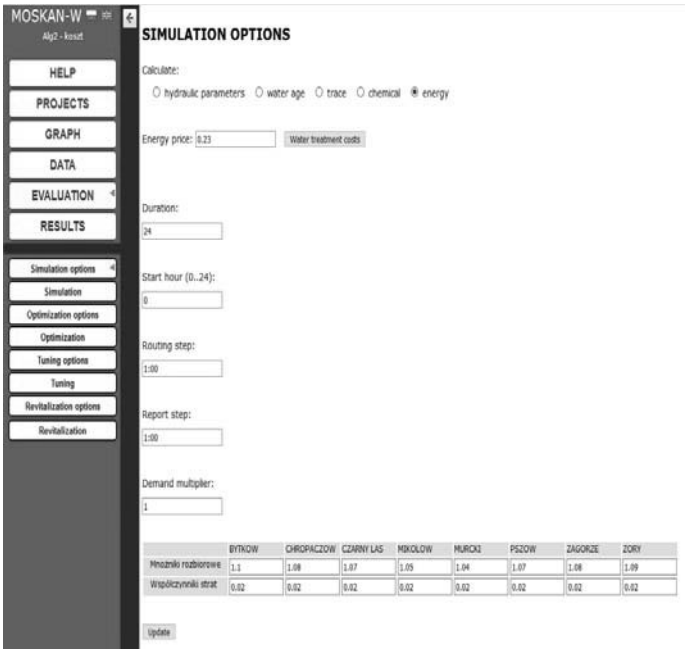


Fig. 8. Starting the algorithm of pumps control in pumping stations of the water supply network.

In the further presented task of pump control, in order to reduce the operating costs of the water supply network, an algorithm is used to minimize the costs of energy consumed by pumps by changing their rotational speed, assuming that these speeds are changed to a small extent to avoid problems with the inaccessibility of the algorithm of hydraulic calculations of the water network model. This task differs from the previous one in that the calculated operating costs additionally include water treatment costs and that the hydraulic calculations are performed for the model of the entire water supply network, so they are global in nature.

Fig. 8 shows the screen of the program used to start calculations concerning the reduction of global energy costs. The data entered at this stage of calculations concern: a given calculation option (energy), a given electricity price (0,23 PLN/kWh), simulation time of network operation (24 h), network load multiplier (1), as well as consumption multipliers and loss coefficients for individual pressure zones of the network. The GPW S.A. network is divided into 8 pressure zones corresponding to eight water intake and treatment stations and for each such zone one should provide coefficients corresponding to individualised water losses resulting from inaccuracy of water meters installed in a given zone (Consumption multipliers) and coefficients corresponding to individualised water losses resulting from network failure in a given zone (Loss coefficients). The values of the loss factors and consumption multipliers are determined on the basis of the operational experience of the operator of each zone. Since in this task, apart from the unit costs of energy consumed by the pumps (0,23 PLN/kWh), the costs of water treatment, different for different treatment plants, are additionally included, it is necessary to load the data on these costs from an external text file. The water treatment costs file contains the price of injecting a cubic metre of water into the network, including the treatment costs and any other costs proportional to the volume of water abstracted. These costs vary from one water intake station to another, so they are assigned to the pump identifiers of the individual water intake pumping stations to enable accurate calculation of overheads. The downloaded water treatment costs from the file enable the change of the costs depending on the seasons and the degree of pollution of the water from the surface water intakes. In order to start optimization calculations, the cost criterion of quality is defined using the logic of fuzzy sets and taking on the triangular function of belonging the criterion value to the fuzzy set. In this way, the variable costs associated with pumping water to the network and its treatment will be minimized (Fig. 9). These costs are approximately proportional to the amount of water pumped into the network by a specific pumping station and are included in the quality criterion on the basis of the coefficients (energy costs + treatment costs) provided by the water supply company, which are different for different plants.

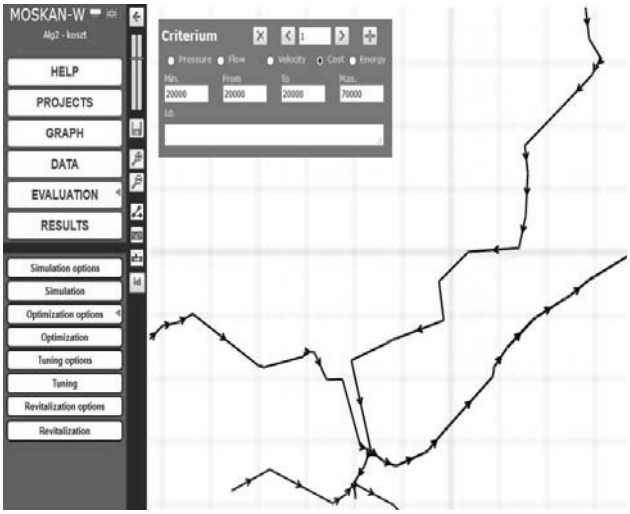


Fig. 9. Definition of a quality criterion for the calculation of global energy cost minimization.

Since cost minimization may lead to unacceptable solutions, for example to the appearance of too low pressures in network nodes, it is possible and recommended to define additional pressure criteria in the task, which will prevent it by limiting the search area accordingly. Control variables taken into account in the algorithm are the speeds of the pumps selected for optimization (Fig. 10).

Fig. 10. Selection of pumps for optimization (selection of control variables) and selection of permissible speed ranges.

Fig. 11 shows the results of cost optimization with regard to control of two pumping stations out of eight that operate on the GPW S.A. water supply network. The results of calculations obtained after 2 and 15 steps of the algorithm's operation are shown. Fig. 12 shows the results for an example where the control of five pumping stations, other than those from the previous task, is taken into account.

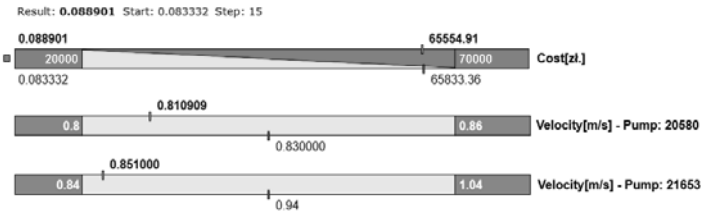


Fig. 11. Results of optimization calculations for two pumps after 15 steps of algorithm operation.

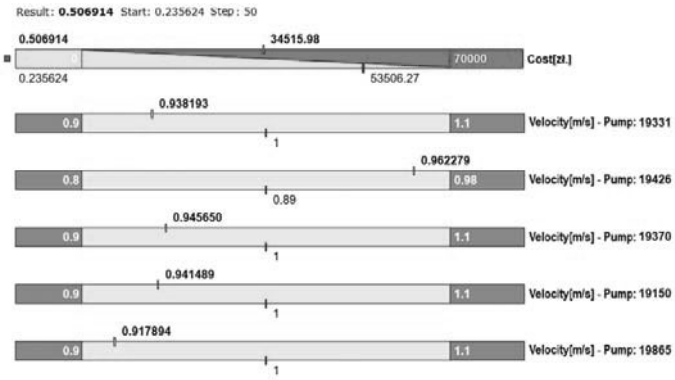


Fig. 12. Results of optimization calculations for five pumps after 50 steps of algorithm operation.

The evaluation of the effectiveness of the presented algorithm of pump control on a global scale is as follows: In the first calculation example, the following results were obtained: pump speed 20580 was reduced by 2% and pump speed 21653 was reduced by 10%. As a result of these reductions, the energy costs associated with the operation of the network were reduced by 0.4%. The obtained result is not satisfactory, therefore in the next calculation example it was decided to take into account 5 pumping stations, which in total consume about 63% of energy connected with water treatment and its pumping into the network. As a result of the new calculations, the following results were obtained: the speed of 4 pumps was reduced by 6%, 5%, 6% and 8% respectively, while the speed of one pump 19426 increased by 8%. At the same time, the energy costs associated with the operation of the network were reduced by 36%.

On the basis of the calculations, the following conclusions can be drawn:

- calculations made in the second example mean that by controlling the speed of several pumps, with the set costs of their operation and the costs of water treatment in the respective pumping station, it is possible to significantly change the load of individual pumping stations, relieving the pumping stations of more expensive ones in favour of cheaper ones.
- relieving a specific pumping station means reducing the speed of the pump running in that pumping station and loading the pumping station means increasing the speed of the relevant pump.
- when starting the algorithm described above, it should be taken into account as control variables of as much as possible, and preferably all water intake and treatment stations.
- the developed algorithm enables a notable reduction of energy consumption in the network.

FINAL REMARKS

The presented MOSKAN-W system is an ICT system dedicated to design, model and optimize water supply networks. It is a tool for rational management of water networks, accomplished with the use of modern information technologies used to build internet applications. The data for MOSKAN-W is loaded from the database of the GIS system. The computing kernel of the application for hydraulic calculations is the EPANET2 software developed by the American Environmental Protection Agency (EPA) and available on the Internet. The original proprietary solution of the application is the interface of MOSKAN-W developed with consideration of its users' convenience and ease of use, as well as optimization algorithms proposed and implemented in the system. In the system the basic group of functions relate to hydraulic and optimization calculations, but apart from the algorithms of optimal speed selection and pump operation control presented in the article, there are also algorithms concernig optimal selection of diameters of water supply pipes, control of the water flow in the network, global optimization of water production costs, calibration of the hydraulic network model, for planning the network revitalization, detecting anomalies and faults in network operation. These algorithms solve the issues that were formulated as a result of consultations with the engineering staff of the GPW S.A. water supply company. The MOSKAN-W program is an open system and can be freely extended by other algorithms solving additional problems related to water supply network management, if they are properly formulated and reported to the authors of the software.

REFERENCES

Fajdek B., Stachura M., Studzinski J.: *Water distribution network optimization using genetic algorithms.* Industrial Simulation Conference ISC'2015 Conference, Valencia, June 1-3, 2015, EUROSIS-ETI 2015, 159-164.

Sluzalec A., Studzinski J., Ziolkowski A.: *MOSKAN-W - the web application for modelling and designing of water supply system*. In: Simulation in Umwelt- und Geowissenschaften, Reihe: Umweltinformatik, ASIM-Mitteilung AM 150, Workshop Osnabrück 2014 (J. Wittmann, Hrsg.), Shaker Verlag, Aachen 2014, 143-153.

Studziński J., Kurowski M.: *Some algorithms suporting the water network management by use of simulation of network hydraulic model*. Industrial Simulation Conference (ISC'2014), 11-13.06.2014, Skoevde Sweden, EUROSIS-ETI 2014, 33-37.

Studzinski J., Wojtowicz P.: *Kalibrierung des hydraulischen Models des Oberschlesischen Wasserversorgungs- und Verteilungsnetzes anhand der Monitoringsdaten aus SCADA und AMR*. In: Modelierung und Simulation von Ökosystemen, Reihe: Umweltinformatik, Workshop Kölpinsee 2015 (Nguyen Xuan Thinh, Hrsg.), Rhombos-Verlag Berlin, 2016.

Wójtowicz P., Pawlak A., Zimoch I., Studziński J.: *Hydraulic modelling and calibration of the complex water distribution system of the Upper Silesian Waterworks PLC.* Konferencja Naukowo-Techniczna: Aktualne zagadnienia w uzdatnianiu i dystrybucji wody, 10-12.06.2015, Szczyrk.

https://www.epa.gov/water-research/epanet

https://www.microimages.com/documentation/Tutorials/Epanet2UserManual.pdf

ANALYZING THE COUPLING PROCESS OF DISTRIBUTED MIXED REAL-VIRTUAL PROTOTYPES

Peter Baumann
Oliver Kotte
Robert Bosch GmbH
peter.baumann5@de.bosch.com
oliver.kotte@de.bosch.com

Lars Mikelsons
Chair for Mechatronics
University of Augsburg
Augsburg, Germany
lars.mikelsons@informatik.uni-augsburg.de

Dieter Schramm
Chair for Mechatronics
University of Duisburg-Essen
Duisburg, Germany
schramm@imech.de

KEYWORDS

Computer Aided Engineering, XiL, Co-Simulation, Real-Time, Time-Delay Compensation

ABSTRACT

The ongoing connection and automation of vehicles leads to a closer interaction of the individual vehicle components, which demands for consideration throughout the entire development process. In the design phase, this is achieved through co-simulation of component models. However, complex co-simulation environments are rarely (re-)used in the verification and validation phases, in which mixed real-virtual prototypes (e.g. Hardware-in-the-Loop) are already available. One reason for this are coupling errors such as time-delays, which inevitably occur in co-simulation of virtual and real-time systems, and which influence system behavior in an unknown and generally detrimental way. This contribution introduces a novel, adaptive method to compensate for constant time-delays in potentially highly nonlinear, spatially distributed mixed real-virtual prototypes, using small feedforward neural networks. Their optimal initialization with respect to defined frequency domain features results from a-priori frequency domain analysis of the entire coupled system, including coupling faults and compensation methods. A linear and a nonlinear example demonstrate the method and emphasize its suitability for nonlinear systems due to online training and adaptation. As the compensation method requires knowledge only of the bandwidths, the proposed method is applicable to distributed mixed real-virtual prototypes in general.

INTRODUCTION

Current trends in the automotive sector like connected and autonomous driving functions are leading to a closer coupling of different vehicle domains. This is for example the case in the development of an emergency brake assistants through the interaction of longitudinal control and brake management. In order to enable a time- and cost-efficient development of such cross-domain vehicle functions, the interactions of the domains must be considered at an early stage of the development process. Simulation experts interconnect the different simulation models, by coupling various tools using the methods of co-simulation [9]. This way complex cross-domain model-in-loop (MiL) co-simulations are implemented. Since the modeling and integration effort to set up such co-simulations is high, the demand is coming up to use the same co-simulation environment not only during the design phase, but also in the verification and validation phase of the development process. In those phases, first components of the vehicle are available as real hardware on test benches. Coupled to the existing MiL co-simulation, detailed mixed real-virtual prototypes are realized to test the hardware or software in open context under realistic (e.g. traffic) conditions. But, due to the real-time requirement, couplings between simulation models and hardware (HiL) always come with coupling faults (e.g. time-delay or measurement noise). If, in addition, the unchanged models from the MiL co-simulation are to be used, the coupling faults even increase, since the models can generally not be compiled on a real-time operating system, but run on a standard Windows PC. Furthermore, there are use cases for cross-company collaboration using mixed real-virtual prototypes, since the complexity of the systems is increasing and their handling requires a wide range of different competencies, which most companies do not have in house. The additional distance between the coupled systems in those use cases further increases the coupling faults and their negative effect on the overall coupled system.

This contribution addresses the time-delays in distributed mixed real-virtual prototypes, by proposing a novel, adaptive compensation method combined with a detailed analysis of the dynamic effects a simulator distribution has on the overall system. After covering related work from the fields of distributed mixed real-virtual prototypes and coupling fault compensation, the compensation method based on a feedforward neural network is presented. The results of the following analysis of the overall system in frequency domain are then used for an optimal compensation method design. Finally, the applicability of the methodology on a nonlinear example is shown.

RELATED WORK

Recently the development of mixed real-virtual prototypes in the automotive sector gained some momentum through the release of the Distributed Co-Simulation Protocol (DCP) [12]. The DCP is designed to standardize the

coupling of real-time or non real-time simulators and thus reducing the integration effort of spatially distributed prototypes.

In literature many examples for the transition from virtual to real testing using mixed real-virtual prototypes can be found. A toolchain for a seamless transition from a MiL co-simulation to heterogeneous HiL testing is presented [11]. Using this toolchain, an engine test bench is coupled to real-time vehicle dynamics and environment models. Mixed real-virtual prototypes can also be used to incorporate the driver in the testing process, e.g. by investigating the interaction between humans and an automatic transmission in a driving simulator [15]. Since test benches usually do not stand side by side, it is reasonable to use the internet when coupling them. Examples for mixed real-virtual prototype coupled via the internet can be found in [7] and [16]. In [2] additionally the DCP is used for the integration.

In all mentioned examples, the focus is on implementing the coupling itself. Coupling faults as delays and dropouts, specified in [16], are known but the use of compensation methods is not yet widespread. In the scientific fields of telerobotics [13] and networked control systems [1] many methods are developed to compensate for these coupling faults, but the compensation is implemented in the controls themselves. For mixed real-virtual prototypes, however, the physical models should not be modified, which is why the compensation must be implemented in the coupling signals [17]. In [18] a model-based-coupling method is presented which is meant for the usage for mixed real-virtual prototypes and is tested on an engine test bench. The parameters of two second order linear systems are identified online to compensate for the delay. A linear fourth order FIR filter is introduced in [17]. Together with a recursive least squares algorithm as identification method, the filter is designed to cope with communication time delays, data-losses and noisy measurements.

However, there is no compensation method capable of representing nonlinear signal behavior. In addition, the influence of the compensation on the overall system should be predictable and verifiable and its parameterization should ideally be based on the dynamics of the coupled system itself.

COMPENSATION METHOD

This paper considers the coupling process between two distributively coupled systems (e.g. vehicle components) "A" and "B", at least one of which is a real-time system. The communication or macro step size ΔT is the constant rate at which data is exchanged between the two systems. The coupling faults that occur in a distributed coupling of mixed real-virtual prototypes are attributable to effects like communication time-delay, jitter, determinism and message loss [16]. Since the time-delay is usually the most dominant fault in a distributed system, the others are neglected in this paper. This simplifies the faults to a constant time-delay τ which represents the time between sending a message from "A" and receiving it at "B" and vice versa. The macro step size fixes its resolution, it holds $\tau = k \cdot \Delta T$ with $k \in \mathbb{N}$.

To compensate for the time-delay, an algorithm is needed which is placed at each input u of each system participating in the distributed real-time co-simulation. The compensation method extrapolates the delayed input $u_{t-\tau}$ to get a predicted input value $\hat{u}_t$ at time t. In order to make the method applicable for as many simulation tools and real-time systems as possible, the extrapolation is signal-based and only the current as well as past values (and no derivatives) of the inputs are used for the extrapolation. Figure 1 gives an overview of the coupled system, the subscript of y, u and $\hat{u}$ stand for the point in time at which the signal is evaluated.

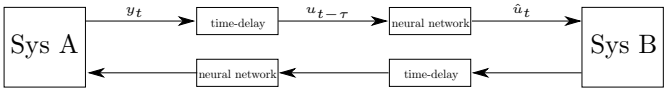


Figure 1: Overview of the Considered Coupled System

For extrapolation, small feedforward neural networks are used here. If the activation function of all network nodes is linear and a vector $\vec{u}$ consisting of p past signal values is the input to the network, the output of the network reads as

$$\hat{u}_t = \vec{a}^T \vec{u} + b \text{ with } \vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} \text{ and } \vec{u} = \begin{bmatrix} u_{t-\tau} \\ u_{t-\tau-\Delta T} \\ \vdots \\ u_{t-\tau-(p-1)\Delta T} \end{bmatrix}. \tag{1}$$

The vector $\vec{a}$ and the scalar b are calculated from the weights of the network. As already stated by [6], this equals a linear autoregressive model, which is a generalization of commonly used extrapolation methods in co-simulation as zero-order-hold (constant extrapolation, ZOH) and first-order-hold (linear extrapolation, FOH). The recursive FIR-Filter for time-delay compensation in [17] is based on the same function.

The main advantage of using a feed forward neural network for the extrapolation is the fact that it can be extended to represent nonlinear signal behavior, by choosing a nonlinear activation function in the hidden layer(s). If Rectified Linear Units (ReLU) or, to prevent parts of the network from "dying", leaky ReLUs are chosen as activation function, the network is enabled to switch between different configurations of the parameters $\vec{a}$ and b. For example, a hidden layer with n neurons with (leaky) ReLU activation implements n^2 different parameter sets of $\vec{a}$ and b depending on which hidden node is active and which is not. But between the switching points the network retains its linear behavior from equation (1), which allows a detailed analysis of the compensation behavior, even in the nonlinear case. Furthermore, the structure of the neural network allows using efficient algorithms for the initial training as well as an online adaption of the weights of the network.

ANALYSIS OF THE COUPLING PROCESS

The coupling faults influence the behavior of a distributed mixed real-virtual prototype significantly. It is not unlikely that an originally stable system gets unstable due to the coupling and the associated coupling faults. In order to increase the confidence in mixed real-virtual prototypes by making statements on robustness regarding the coupling faults, this section aims to analyze the entire closed-loop system including coupling faults and compensation method in frequency domain. First, the transfer function of the coupling process including coupling faults and compensation method is calculated and second it is shown using an example how this transfer function can be utilized to analyze the influence of the coupling faults and of the compensation method on the closed-loop system. Furthermore, it is shown in the next section that the compensation method can be initialized optimally based on this analysis.

As already mentioned in [3], a detailed analysis of a co-simulation is also possible in time domain via a multi-rate approach, but due to the variable step size solvers usually used for physical models, the analysis would become very complex. Instead, in this paper, the analysis is carried out in frequency domain which leads to the following assumptions:

1. All sub-systems, their inputs and their outputs are assumed to be ideally time continuous. It follows that the numerical errors made by a solver due to evaluating the system equations at discrete points in time are neglected. The measurement error that occurs when reading out values of real-time systems (e.g. test benches) is also neglected.
2. The macro step size ΔT is constant and small enough to avoid aliasing in all coupling signals. This is verified via the Nyquist-Shannon theorem with the condition $\bar{\omega}\Delta T << \pi$ for the ratio of the maximal bandwidth $\bar{\omega}$ of the coupling signal and ΔT [4].
3. The coupling faults are simplified to a constant time-delay τ (see previous section).

Since all inputs and outputs of the sub-systems are time continuous, it is reasonable to assume that also the coupling process is a time continuous element which includes the overall correlation between the continuous output signal of a sub-system and the continuous input signal of another sub-system. The coupling process includes two different effects. On the one hand, the disturbing effects of sampling and delay due to data exchange and communication and, on the other hand, the added methods of compensation and reconstruction to compensate for these effects. Figure 2 displays an overview of two distributively coupled systems "A" and "B" and the coupling element.

To enable a closed-loop analysis of the coupled system in frequency domain, the Laplace transform is applied to each component separately. The transfer function of the disturbing effects $G_f(s)$ describes the correlation between an output signal $y(t)$ and the ideally reconstructed input $u_s(t)$ of another sub-system in frequency domain. Hence,

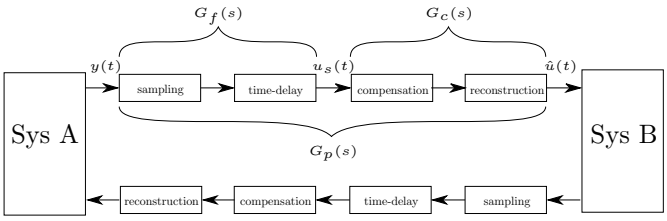


Figure 2: Components of the Real-Virtual Prototype

under assumption 2) and 3) holds

$$U(s) = \underbrace{\frac{1}{\Delta T} \cdot e^{-s\tau}}_{G_f(s)} \cdot Y(s). \tag{2}$$

A detailed derivation on why the influence of the sampling can be represented by a scaling with $\frac{1}{\Delta T}$ under these circumstances can be found in [4]. The authors carry out the Laplace transformation of a coupling process of a non-iterative offline co-simulation (without delay and compensation).

The transfer function $G_c(s)$ represents the frequency domain correlation of the ideally reconstructed input $u_s(t)$ and the applied input $\hat{u}(t)$. It holds

$$\hat{U}(s) = G_c(s) \cdot U(s). \tag{3}$$

$G_c(s)$ is now derived for a feedforward neural network with linear activation function as compensation and ZOH as the reconstruction method. Starting point to get $G_c(s)$ is equation (1) which contains the sampled behavior of the compensation method under consideration of the constant time-delay τ. Including the ZOH reconstruction leads to a piece wise constant function in time domain

$$\hat{u}(t) = \vec{a}^T \vec{u} + b \quad \text{with} \quad n\Delta T \leq t < (n+1)\Delta T. \tag{4}$$

The Laplace transform of $\hat{u}(t)$ is defined as

$$\mathcal{L}\{\hat{u}(t)\}(s) = \sum_{n=0}^{\infty} \int_{n\Delta T}^{(n+1)\Delta T} \hat{u}(t) e^{-s\Delta T} dt. \tag{5}$$

The linearity property of the Laplace transform allows a separate transformation of each summand of $\hat{u}(t)$. For the first summand $\hat{u}_1(t) = a_1 \cdot u_{t-\tau}$ equation (5) simplifies to

$$\mathcal{L}\{\hat{u}_1(t)\}(s) = \hat{U}_1(s) = \sum_{n=0}^{\infty} \int_{n\Delta T}^{(n+1)\Delta T} \hat{u}_1(t) e^{-s\Delta T} dt. \tag{6}$$

Since $\hat{u}_1(t)$ is piece wise constant in, it holds

$$\hat{U}_1(s) = \sum_{n=0}^{\infty} \hat{u}_1(t) \frac{e^{-sn\Delta T} - e^{-s(n+1)\Delta T}}{s} \tag{7}$$

$$= \underbrace{a_1 \frac{1 - e^{-s\Delta T}}{s}}_{G_{c1}(s)} \underbrace{\sum_{n=0}^{\infty} u_{t-\tau} e^{-sn\Delta T}}_{U_1(s)}. \tag{8}$$

This results in the Laplace transform of the first summand $G_{c1}(s)$ of $G_c(s)$, which is equal to a ZOH extrapolation

scaled by a_1. The transform for all other summands of $\hat{u}(t)$ works accordingly under consideration of an additional time shift. Therefore it holds

$$G_c(s) = \sum_{n=0}^{p-1} a_{n+1} e^{-ns\Delta T} \cdot \frac{1 - e^{-s\Delta T}}{s} + \frac{b}{s} e^{-s\Delta T} \quad (9)$$

for a compensation taking p past signal values into account. Finally, using equations (2) and (3), the overall coupling process in frequency domain $G_e(s)$ can be written as

$$G_p(s) = G_f(s) \cdot G_c(s) \quad (10)$$

$$= \sum_{n=0}^{p-1} a_{n+1} e^{-(\tau+n\Delta T)s} \cdot \frac{1 - e^{-s\Delta T}}{s\Delta T} + \frac{b}{s\Delta T} e^{-(\tau+\Delta T)s}. \quad (11)$$

Remarks:

- The methodology also allows using other methods to reconstruct the input between macro time steps (e.g. FOH) by changing equation (4) accordingly [4].
- When using a piece wise linear activation function in the neural network (e.g. ReLU or leaky ReLU) the behavior is nonlinear. Of course in this case the Laplace transform is not valid any more, but since the network will still maintain a piece wise linear structure, each linear component can be transformed separately, to examine the different possibilities of the transmission behavior of the neural network.

Linear Two-Mass Oscillator

Using the example of a two-mass oscillator, it is shown how the transfer function of the coupling process $G_p(s)$ can be used to estimate the overall system behavior of a distributively coupled system. Here, this is done purely in simulation with synthetic coupling faults between the sub-systems.
Two via spring and damper connected models of single mass oscillators form the two-mass oscillator as can be seen in figure 3. At the dashed line, the oscillator is divided into two sub-models each containing one mass. The models are coupled using the so called force/displacement coupling [14] approach, where one model calculates the coupling force and receives position and velocity of the mass of the other model. Table 1 contains the parameters of the slightly damped coupled system.

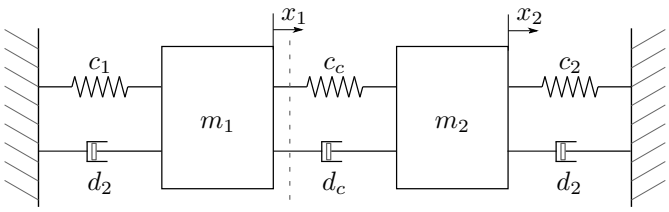


Figure 3: Linear Two-Mass Oscillator

A Laplace transform of the system equations of the two single mass oscillators [14] yields their transfer functions $G_{mass1}(s)$ and $G_{mass2}(s)$. Together with the transfer function of the coupling process $G_p(s)$ the coupled system can be interpreted as a control circuit, which allows stability analysis with Bode and Nyquist plots. Figure 4 sketches the idea.

Table 1: Parameters of the Two-Mass Oscillator (SI)

Parameter	Value
$m_1, m2$	$100, 1$
c_1, c_2, c_c	10
d_1, d_2, d_c	0.01

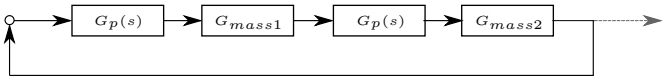


Figure 4: Interpretation of the System as Control Circuit.

In order to determine $G_p(s)$ explicitly, numerical values must be assigned to the parameters of the distributed simulation and the compensation method. The macro step size is set to $\Delta T = 0.001s$ and the delay per coupling direction to $\tau = 0.003s$. The resulting round-trip-time of $RTT = 0.006s$ is much lower than e.g. in a coupling via the internet, but since the two-mass oscillator is a strongly coupled system, even this short round-trip-time has an effect on the system behavior.
The neural network which is used for compensating the delay is implemented in Python using the Keras package [5]. The size of the network and especially the number of inputs (considered past signal values) is a trade-off between computation resources needed and the capability of the network to represent nonlinear signal behavior. With a linear activation function of course, the capability of the network does not increase with the number of neurons [6], but for the time-delay compensation of a nonlinear signal a single neuron is not sufficient. The chosen network size of four input neurons, two hidden neurons and one output neuron is the smallest, which gives good results also for nonlinear signals (see nonlinear example).
For this example, the weights of the network are obtained by training the network before simulation on a self-created data set which is based on simulation results of the two-mass oscillator with different initial conditions (without faults). Since this is cumbersome, section presents how the information from the analysis shown here can be used to optimally parameterize the network in advance without training. After the training the parameters (equation (1)) calculated from the weights of the network read as

$$\vec{a} = \begin{bmatrix} 2.4748 \\ -0.6470 \\ -0.1664 \\ -0.6664 \end{bmatrix}, b = 0. \quad (12)$$

Thus, the coupling process including compensation for this example can be determined by equation (11). Multi-

plying the transfer functions, as shown in figure 4, results in the open-loop transfer function representation of the overall distributively coupled system

$$G_{sys}(s) = G_{mass2}(s) \cdot G_p(s) \cdot G_{mass1}(s) \cdot G_p(s). \quad (13)$$

Figure 5 shows the open loop Bode plot of $G_{sys}(s)$. The blue curve is the reference without coupling effects ($G_p(s) = 1$), red curve is with included faults and trivial (ZOH) compensation and the yellow curve is with included faults and neural network compensation. First of all, the two resonance frequencies of the two-mass oscillator are clearly visible and, more importantly, they do not change by adding faults or the compensation method. This means the qualitative system behavior stays the same. The Bode plot is also useful to verify, that the frequency bandwidth of the system (areas with high frequency amplification) is below 6rad/s and therefore small enough to avoid aliasing (assumption 2). It is further observed, that the magnitude for high frequencies is larger when the compensation method is included in the control circuit. However, this amplification is not critical, since the two-mass oscillator itself dampens very strongly in this frequency range. In the bode phase diagram a rapidly decreasing phase for high frequencies is visible, which is typical for systems with delays. It is noticeable that without compensation the phase deviates away from the reference within the bandwidth, whereas with compensation the phase follows the reference in a larger frequency range.

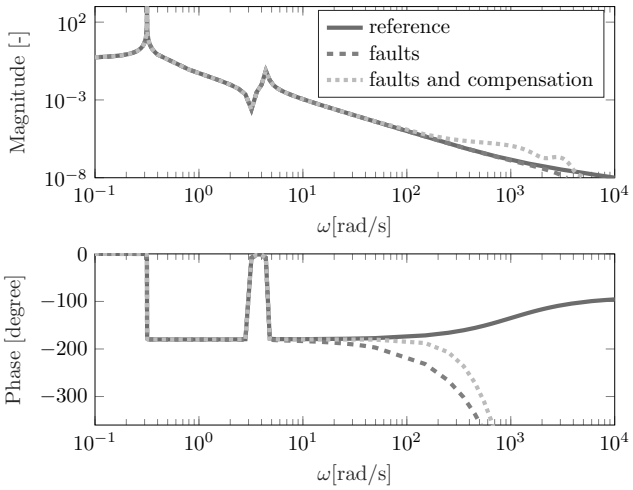


Figure 5: Bode Plot of $G_{sys}(s)$

How this affects stability can be investigated using the Nyquist plot in figure 6. A close look at the area around the critical point $P_c = (-1, 0j)$ reveals a deviation of the Nyquist locus in the case without compensation. The encirclement of the Nyquist locus around the critical point is different from the reference when the faults are introduced and changes back to the reference when the compensation is added (Nyquist Stability Criterion). This behavior is confirmed with simulation results: The system with faults is unstable without the compensation and can be stabilized by adding the compensation method. The reference system is of course stable, since a two-mass oscillator is a mechanically stable system.

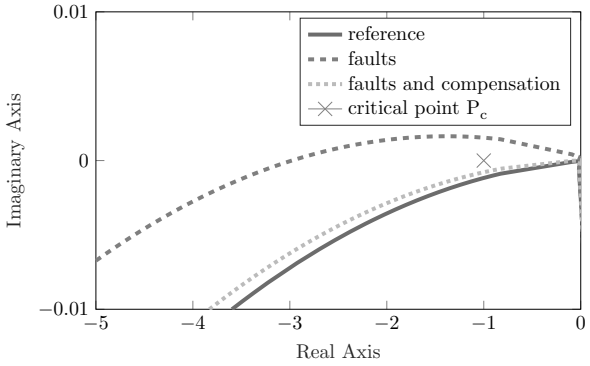


Figure 6: Enlargement of Nyquist Plot of $G_{sys}(s)$ with a pre-trained Neural Network as Compensation Method

This section revealed two important things: Firstly, an analysis of the overall system behavior of a distributed real-virtual prototypes in the frequency domain is possible. This allows, for example, to check in advance whether a system can be stabilized under a certain time-delay with a certain compensation and reconstruction method. Of course these statements are subject to some simplifications and assumptions and can therefore not be transferred one-to-one to reality, but the basic system behavior under certain coupling errors can be demonstrated with this analysis approach. Secondly, it is shown that the trained neural network is able to compensate for the coupling faults.

DESIGN OF THE COMPENSATION METHOD

The idea is to use the analysis method from the last section not only for checking the system behavior under the influence of delays and a compensation method, but also for an optimal design of the compensation method itself. This eliminates the need for training the neural network in beforehand.

The compensation method is very flexible due to the parameters $\vec{a}$ and b. Therefore, requirements for the behavior of the compensation in the frequency domain are made first, before on that basis an optimization problem is defined, by whose solution the optimal parameters are found. Since the compensation depends strongly on the configuration and properties of the coupled systems, the requirements are defined for the transfer function of the overall coupling process $G_p(s, \vec{a}, b)$.

The parameters $\vec{a}$ and b must be chosen such that

- $\sum_p a_p + b = 1$, which leads to $\lim_{w \to 0} G_p(s = j\omega) = 1$ and thus a correct extrapolation of constant signals.
- the magnitude of $G_p(s)$ within the bandwidth of the coupled system is neutral ($|G_p(j\omega)| = 1$).
- the phase shift of $G_p(s)$ within the bandwidth of the coupled system is neutral ($\angle G_p(j\omega) = 0°$).
- outside the bandwidth frequencies of the coupled system, the combined magnitude of all coupling pro-

cesses $G_p(s)$ does not increase faster than the magnitude of the coupled system decreases. This guarantees that additional amplifications of $G_{sys}(s)$, which are introduced by $G_p(s)$, will be damped by the dynamics of the coupled system itself.

The second and third requirement ensure the compensation of the coupling faults in frequency ranges where the coupled system is dynamically active.
All requirements are weighted and combined into a single objective function, but a multi-criteria optimization would also be possible. The combined objective function of the optimization problem with $s = j\omega$ reads as

$$J(a,b) = \alpha J_a + \beta J_p + \gamma J_r, \text{ with} \tag{14}$$

$$J_a = \int_{\omega_{bw,min}}^{\omega_{bw,max}} \frac{1 - |G_p(j\omega)|}{\omega_{bw,max} - \omega_{bw,min}} d\omega \tag{15}$$

$$J_p = \int_{\omega_{bw,min}}^{\omega_{bw,max}} \frac{\angle G_p(j\omega)}{\omega_{bw,max} - \omega_{bw,min}} d\omega \tag{16}$$

$$J_r = \int_0^{\omega_{bw,min}} \max\left(|G_p(j\omega), 1\right) d\omega - \omega_{bw,min} \tag{17}$$

$$+ \int_{\omega_{bw,max}}^{\frac{2\pi}{\Delta T}} \max\left(|G_p(j\omega)| - \left(\frac{\omega}{\omega_{bw,max}}\right)^v, 0\right) d\omega. \tag{18}$$

The bandwidth frequencies of the coupled system are within the interval $[\omega_{bw,min}, \omega_{bw,max}]$ and the remaining frequencies (up to the sampling frequency) in $[0, \omega_{bw,min})$ and $(\omega_{bw,max}, \frac{2\pi}{\Delta T}]$. The optimization problem results in

$$\min_{a,b} J(a,b) \text{ such that } \sum_p a_p + b = 1. \tag{19}$$

Remarks:

1. It shall apply $\gamma >> \alpha, \beta$, to ensure that the coupling process stays inside the magnitude boundary for frequencies outside the bandwidth frequencies. Furthermore holds $\alpha = 100\beta$ to punish a phase difference of 1° equally as a magnitude error of 1% [4].
2. Parameter v depends on the relative degree r of the coupled system. It holds $v = \frac{1}{2r}$
3. The calculation of the objective function is possible with very little system information. Transfer functions of the coupled subsystems are not necessary. The bandwidth of the coupling signals could instead be estimated by Fourier transformations of the coupling signals and the relative degree can be set conservatively to one in case of doubt.

The bandwidth of the two mass oscillator is in the range $[1\frac{rad}{s}, 6\frac{rad}{s}]$ and the relative degree is $r = 2$. Thus, the numerically found solution of equation (19) is

$$\vec{a}_{opt} = \begin{bmatrix} 6.5103 \\ -1.5509 \\ -9.9296 \\ 5.9702 \end{bmatrix}, b_{opt} = 0. \tag{20}$$

Figure 7 and 8 show the bode plot and the enlarged nyquist plot with the neural network compensation, initialized with the optimized $\vec{a}_{opt}$ and b_{opt}. Similar results as in the trained version of the neural network are obtained. The deviation of the nyquist curve between the reference and the compensated version is even smaller than with the trained neural network.

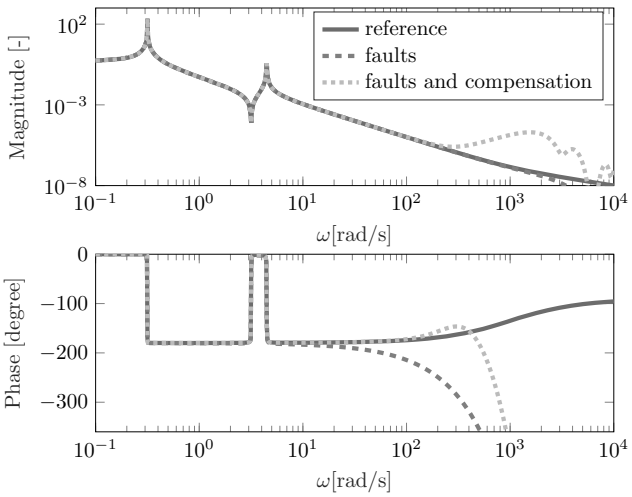


Figure 7: Bode Plot of G_{sys} with an Optimally Initialized Neural Network as Compensation

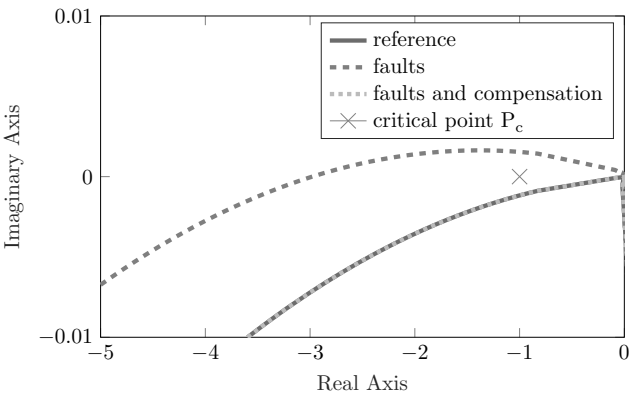


Figure 8: Enlargement of Nyquist Plot of G_{sys} with an Optimally Initialized Neural Network as Compensation

NONLINEAR EXAMPLE

In the last section, an optimal time-delay compensation of the form of equation (1) was designed on the basis of very little system information. Now the question arises why neural networks should be used as compensation method at all, since they are more complex to calculate and implement than equation (1) (and with linear activation function they behave exactly alike). The advantages of the neural network is that it can be adapted online by training in parallel and is also capable of represent nonlinear signal behavior.
To show this, a nonlinearity extends the two-mass oscillator system from section . A mechanical stop prevents the first mass from positioning beyond $x_{1,stopp} = -0.1m$. This is implemented in Matlab/Simulink by reversing the velocity of the first mass $v' = -e \cdot v$ at $x_{1,stopp}$ [8]. The

coefficient of restitution is $e = 0.7$, which leads to a partly inelastic collision. For $x_1 > x_{1,stopp}$ the system behaves linear and corresponds exactly to the two-mass oscillator from section (all system parameters are identical), whereby the optimized compensation parameters from equation (20) are also optimal in the nonlinear system.
Figure 9 shows simulation results with the nonlinear two-mass oscillator. The oscillation starts from the initial condition $x_{1,0} = x_{2,0} = 1$ and no external forces act on the system. The bouncing of the first mass is visible and all curves lie on top of each other, which means that the coupling faults have no significant effect on x_1 during the first $50s$ of the simulation.

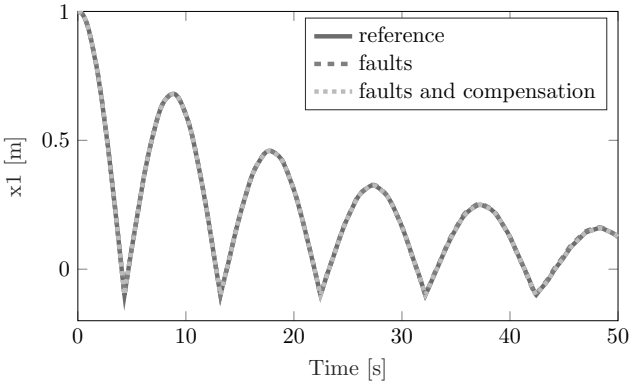


Figure 9: Position of the First Mass for $t \in [0s, 50s]$

Figure 10, which shows the same simulations $400s$ later, reveals a large deviation of the simulation results and also different convergence properties of the coupled system. The slight damping reduces the oscillation amplitude of

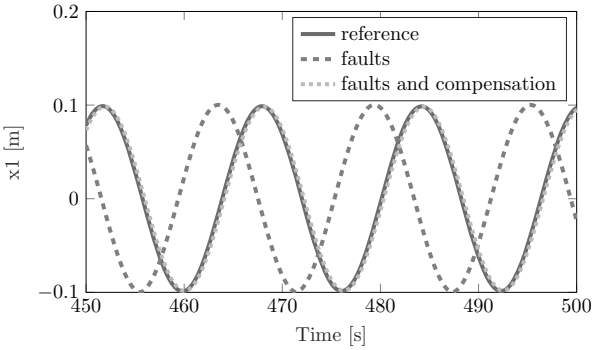


Figure 10: Position of the First Mass for $t \in [450s, 500s]$

x_1 in each period in the reference simulation as well as in the simulation with faults and compensation. In the end all the energy will be dissipated and the system will end up in its equilibrium. In the simulation with faults and without compensation, the oscillation amplitude of x_1 stays at $0.1m$ and the mass hits the mechanical stop in each period. The energy fed into the system each period due to the instability caused by the faults (figure 6) is dissipated by the partly elastic collision with the mechanical stop.

Figure 11 shows, however, the problem of the compensation method designed on the linear system. The velocity reversal at the mechanical stop causes jumps in the coupling signal $\dot{x}_1$, which leads to an overshoot by factor $a_{1,opt} = 6.5103$ of the compensation. In this particular case, the large compensation errors at the jumps do not effect the overall system behavior significantly (system is still stable, see figure 10), but in general such large errors should be avoided.

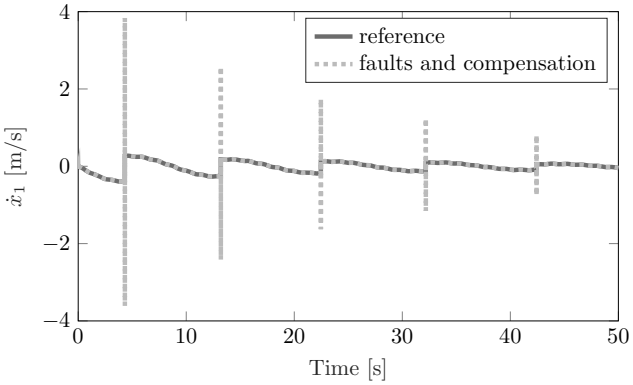


Figure 11: Velocity of the First Mass with Linear Neural Network as Compensation

Here the advantages of the neural network can be fully utilized. Instead of a linear activation function, leaky ReLUs are now used in the two neurons of the hidden layer, which enables the network to switch between four different behaviors. The network is initialized the same way as the linear neural network from the previous example, but is now adapted during the simulation. The online training is performed in parallel in an external process, in order not to influence the computation time of the compensation. For this purpose, the values of the sampled coupling signal are stored during the simulation and sent to the training process together with the current configuration of the neural network. There, training data is created from the data points. Each training sample consist of an input vector $\vec{x}$ with p consecutive data points and a response y, which contains the data point $k = \frac{\tau}{\Delta T}$ steps after the latest value of the input vector. The neural network from the compensation process is then duplicated and optimized in the training process, to minimize the cost

$$C(w) = \frac{1}{N} \sum_{i=1}^{N} L(f(x_i, y_i, w)) \tag{21}$$

for N training samples. The function f stands for the neural network with the weights w. The loss function L implements the mean-squared-error algorithm and the optimizer Adam [10], a gradient-descent algorithm with adaptive learning rate, solves the optimization problem. Once the optimization is complete, the updated weights are sent to the compensation process, where the neural network is updated. This online adaption process can be repeated several times during a simulation.
Figure 12 shows the simulation results of $\dot{x}_1$ again, but this time the online training was active. After the first

jump at 4.5s, the first online training cycle starts and finishes a few seconds later. This way, the network is able to switch its behavior at the second jump and the predicted signal does not overshoot anymore. If the simulation would now be carried out using the adapted neural network as compensation method right from the start, also no overshoot would occur at the first jump.

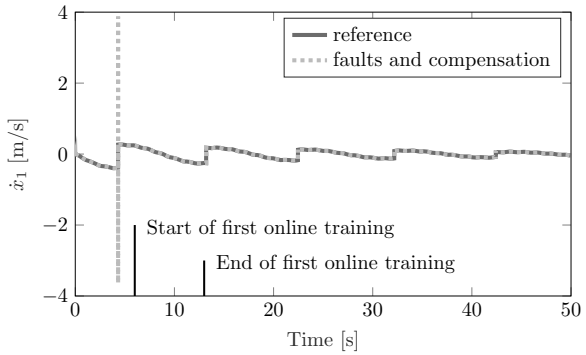


Figure 12: Velocity of the First Mass with Nonlinear and Online Adapting Neural Network as Compensation

CONCLUSION AND FUTURE WORK

Current trends in the automotive sector require distributed mixed real-virtual test approaches. In this work a feedforward neural network is used as a generalized compensation method for the coupling faults (constant time-delay) of mixed real-virtual prototypes. In order to be able to make statements in advance on how the compensation method will behave in a spatially distributed co-simulation, an analysis method in frequency domain is proposed that describes the overall coupling process. Even with very little information on the coupled subsystems, it is possible to use the analysis to determine the parameters of the neural network via optimization in such a way that the coupling faults are compensated optimally. Furthermore, it is shown that the neural network based compensation is also able to adapt to nonlinear signal behavior using an online learning strategy.
Future work will on the one hand focus on the consideration of different fault effects to be able to analyze the coupling process in a more realistic way. On the other hand the entire methodology will be tested on a real, industrial test bench coupled to a MiL simulation.

REFERENCES

[1] John Baillieul and Panos J. Antsaklis. Control and communication challenges in networked real-time systems. *Proceedings of the IEEE*, 95(1):9–28, 2007.

[2] Peter Baumann, Martin Krammer, Mario Driussi, Lars Mikelsons, Josef Zehetner, Werner Mair, and Dieter Schramm. Using the distributed co-simulation protocol for a mixed real-virtual prototype. In *Proceedings of 2019 IEEE International Conference on Mechatronics*, Ilmenau, Germany, 2019. IEEE Industrial Electronics Society.

[3] Martin Benedikt and Edo Drenth. Relaxing stiff system integration by smoothing techniques for non-iterative co-simulation. In *IUTAM Symposium on Solver-Coupling and Co-Simulation*, pages 1–25. Springer, 2019.

[4] Martin Benedikt, Daniel Watzenig, and Anton Hofer. Modelling and analysis of the non-iterative coupling process for co-simulation. *Mathematical and Computer Modelling of Dynamical Systems*, 19(5):451–470, 2013.

[5] François Chollet et al. Keras. `https://keras.io`, 2015.

[6] Georg Dorffner, editor. *Neural networks for time series processing*. Citeseer, 1996.

[7] Tulga Ersal, Mark Brudnak, Ashwin Salvi, Jeffrey L. Stein, Zoran Filipi, and Hosam K. Fathy. Development of an internet-distributed hardware-in-the-loop simulation platform for an automotive application. American Society of Mechanical Engineers, 2009.

[8] Heinz A. Gall. Zur simulation von haftreibung und mechanischen anschlägen. *ASIM Nachrichten*, 1:4–9, 2001.

[9] Cláudio Gomes, Casper Thule, David Broman, Peter Gorm Larsen, and Hans Vangheluwe. Co-simulation: a survey. *ACM Computing Surveys (CSUR)*, 51(3):49, 2018.

[10] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[11] Serge Klein, Rene Savelsberg, Feihong Xia, Daniel Guse, Jakob Andert, Torsten Blochwitz, Claudia Bellanger, Stefan Walter, Steffen Beringer, Janek Jochheim, and Nicolas Amringer. Engine in the loop: Closed loop test bench control with real-time simulation. *SAE International Journal of Commercial Vehicles*, 2017.

[12] Martin Krammer, Martin Benedikt, Torsten Blochwitz, Khaled Alekeish, Nicolas Amringer, Christian Kater, Stefan Materne, Roberto Ruvalcaba, Klaus Schuch, Josef Zehetner, Micha Damm-Norwig, Viktor Schreiber, Natarajan Nagarajan, Isidro Corral, Tommy Sparber, Serge Klein, and Jakob Andert. The Distributed Co-Simulation Protocol for the Integration of Real-Time Systems and Simulation Environments. 2018.

[13] Dale A. Lawrence. Stability and transparency in bilateral teleoperation. *IEEE transactions on robotics and automation*, 9(5):624–637, 1993.

[14] Pu Li. *On the Numerical Stability of Co-Simulation Methods*. PhD thesis, Technische Universität Darmstadt, 2017.

[15] Niko Maas, Benjamin Hesse, Martin Koppers, and Dieter Schramm. Simulator setup according to use case scenarios-a human-oriented method for virtual development. IEEE, 2014.

[16] Viktor Schreiber, Valentin Ivanov, Klaus Augsburg, Matti Noack, Barys Shyrokau, Corina Sandu, and Pieter Schalk Els. Shared and distributed x-in-the-loop tests for automotive systems: Feasibility study. *IEEE Access*, 6:4017–4026, 2018.

[17] Georg Stettinger, Martin Benedikt, Markus Tranninger, Martin Horn, and Josef Zehetner, editors. *Recursive FIR-Filter design for fault-tolerant real-time co-simulation*, 2017.

[18] Georg Stettinger, Martin Horn, Martin Benedikt, and Josef Zehetner. A model-based approach for prediction-based interconnection of dynamic systems. IEEE, 2014.

VIRTUAL REALITY FOR A REALISTIC SIMULATION WITH DYNAMIC MOTION PLATFORMS

E. Thöndel
Department of Electric Drives and Traction
Czech Technical University
166 27, Prague,
Czech Republic
E-mail: thondee@fel.cvut.cz

KEYWORDS
Virtual reality, Simulator, Motion Platform, Optimal Control, Motion Cueing, VR Sickness.

ABSTRACT

The primary purpose of the project is research in the field of virtual reality to be used in general vehicle simulators (including flight simulators or any self-propelled or self-moving equipment). The project is focusing on the development of a complex solution of all scenes VR applications using VR headsets. This fast-growing technology, which can significantly enhance the simulation fidelity, has a high potential for a future wide range of applications, but there are also unsolved problems, which are related to particular applications. One of the unwanted phenomena of virtual reality is simulation-sickness. The modern VR systems require a full range of cues, and one the most important is the motion one. A motion platform with advanced motion cueing algorithm must be used to avoid or at least to minimize the occurrence of VR sickness. This should be done with the use of hi-precision control but also advanced knowledge-based and predictive algorithms. The use of VR headsets also requires low latency and precise tracking of the observer's head in a reference frame of the virtual and real-world assets. Solving the complex of problems the virtual reality deployment recently faces will open a wide range of applications in many fields. In this project, we are focusing on R&D applications, for professional simulators usable in industrial development, education, or medical treatment application. The paper deals with some aspects of deployment of VR in these main areas with a particular focus on general usability and VR-sickness.

INTRODUCTION

Driving and flight simulators are traditionally and historically used mainly in the field of training. In industry, they are used for research and development of new products, prototype, and testing. There is a growing demand for simulators with the use of virtual reality (VR) that has reached the technical level for industrial use in the last few years. The idea of a virtual reality vehicle simulator entirely fits into Industry 4.0's concept of replacing physical prototypes with virtual models that can be easily modified and optimized. If a person is also a part of the device, VR technology is the ideal way to optimize human-machine interaction at the virtual prototype level as well. The application capabilities of the simulation in connection with VR are extensive, but some issues need to be resolved in order for VR technology to exceed the current limit of use in the gaming industry and to meet expectations of Industry 4.0.

The current state of development of vehicle simulators using VR is still at the stage of experiments, and only in the field of driving simulators. Flight simulators or simulators of other means of transport do not use VR technology yet.

A VR driving simulator for HMI testing was developed (2015) in the Centre of Competence HMI, Robert Bosch GmbH. The simulator was tested on a sample of differently experienced drivers, showing the following results; 76 % of the users felt present in VR, and only 20 % still paid attention to the real environment, so the level of fidelity and feeling of presence in the simulated environment is very high. On the other hand, 84 % experienced a feeling of nausea, especially when driving through sharp curves (Ihemedu-Steinke 2015). The main reason for this undesirable phenomenon, which limits widespread use, is the absence of precisely synchronized reproduction of motion effects. If this phenomenon could be suppressed, it would be possible to fully exploit the potential of VR simulators to create high fidelity scenarios to simulate critical situations without endangering the driver's life (Kemeny 2014).

Although VR is not a new technology, its application is associated with many limitations that must be solved to master a VR simulator successfully. Further development needs to focus on the precise reproduction of motion effects, haptic feedback, and high image resolution in VR (Meixner 2017). The latter task is already partially solved by the technical level of VR headsets, although the level of resolution is still not at the optimum level. Haptic feedback can be solved by a mixed reality where the virtual and real-world exactly overlaps. Consequently, follow-up development must be focused on accurate reproduction of motion effects, especially for suppressing nausea resulting from the conflict between the expected and actual response. The high level of feeling of presence in VR also increases the sensitivity of a person to these conflicts, and thus puts increased demands on the precise reproduction of all sensations.

VR VS. CLASSICAL SIMULATOR WITH A MOTION PLATFORM

Exact synchronization of reproduction of individual sensations can only be achieved by deep mutual integration of individual control system components that function more or less independently (a mathematical model of a vehicle, terrain

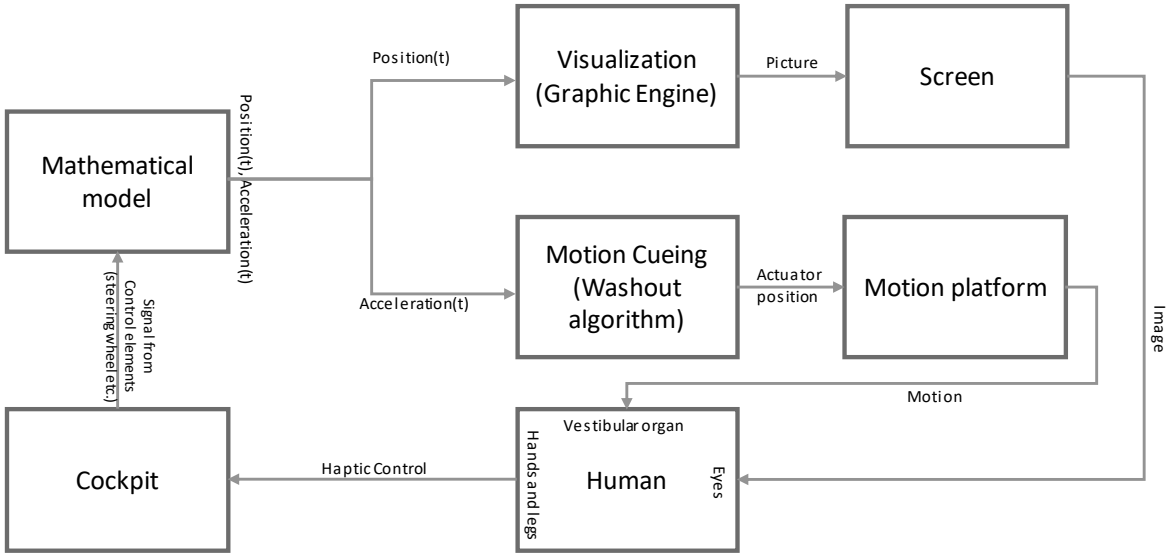


Fig. 1: Standard Logic Structure of a Simulator.

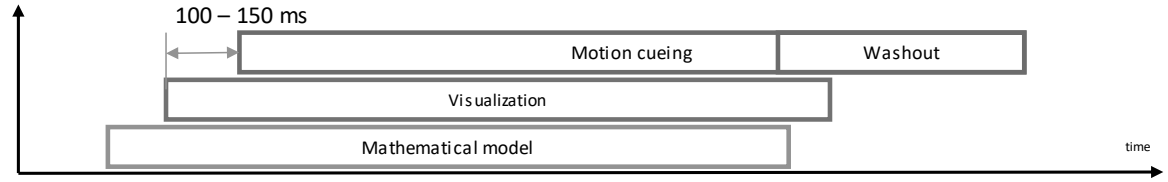


Fig. 2: Time Sequence of Reproduction of Visual and Motion Sensations in a Standard Simulator.

database, motion cueing algorithm and audio-visual engine). In particular, motion effects (acceleration) are phase-shifted in front of the visual effects (position and velocity), but the reproduction is mostly inverted in standard simulators (a motion platform responds with a delay of about 100 – 200 ms). This mismatch or total absence of a motion platform causes the mentioned problems with a rapid onset of feeling sick. The standard logic structure of a simulator is shown in Fig. 1 and Fig 2.

Deep integration means using all available information from all subsystems to predict behavior, in particular, to enable the motion platform to prepare itself in time for the required maneuver. An introduction to the idea of using prediction in motion cueing algorithm is given in the author's article (Thöndel 2013), where the Model Predictive Control (MPC) algorithm is used to detect kinematic limits. The current commercially available VR headsets already implement prediction of head position to refine image synchronization. The prediction in the field of motion simulation is complicated because of the longer prediction horizon (about 150 ms). The logic structure of a VR simulator is shown in Fig. 3, where new systems and links are highlighted. An important new feature is compensation of motion generated by the motion platform to avoid interaction with sensors inside the VR headset. This feature is referred to as the "Motion Compensation" in Fig. 3, and exact synchronization of the motion cueing algorithm and sensors inside the VR headset is required for implementation.

MOTION COMPENSATION

Motion Compensation (Cancellation) is a method of compensating the impact of a motion platform on sensors inside a VR headset. Implementing this compensation is an essential part of every VR simulator equipped with a motion platform.

The principle of this method consists in calculating the inverse movement of the motion platform and adding the resulting position to the position measured by the sensors in the VR headset (Fig. 5). The calculation of the inverse movement is not complicated – it is based on the inversion of the transformation matrix that is used to calculate the position of the motion platform.

A fundamental problem with the implementation of the compensation is the exact synchronization to achieve the time sequence, as shown in Fig. 4. The entire control system is distributed on several independent hardware platforms, so it is necessary to use a communication bus with a guarantee of real-time transmission (e.g., POWERLINK).

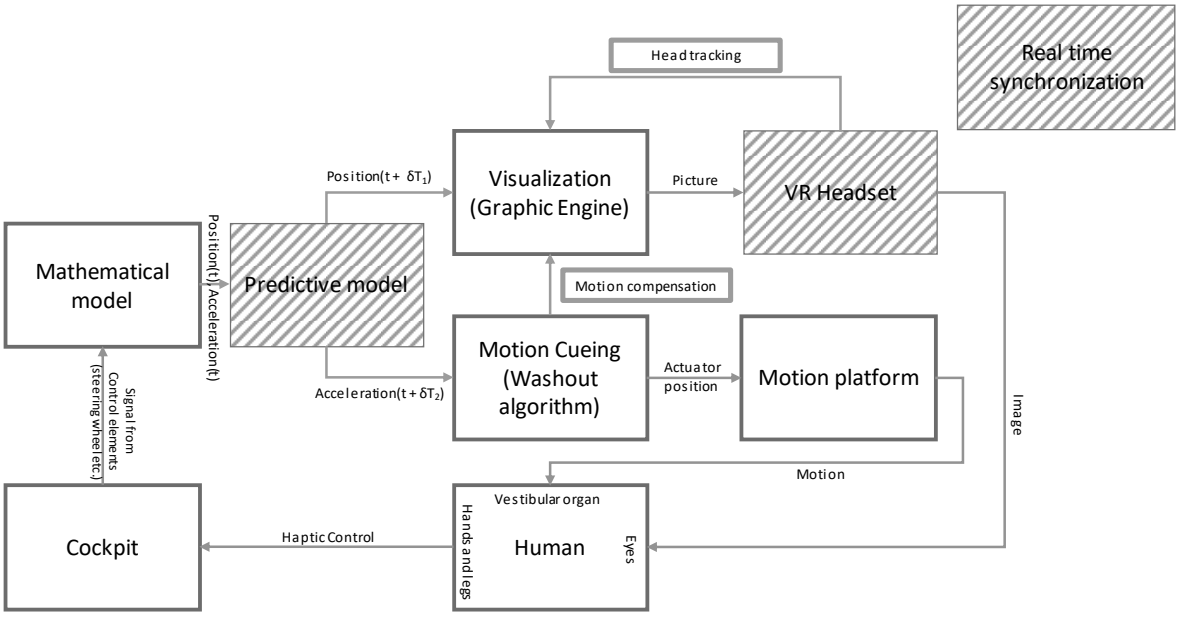


Fig. 3: Logic structure of a VR simulator.

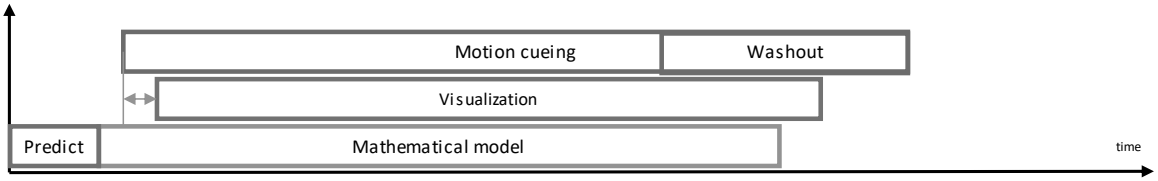


Fig. 4: Time sequence of reproduction of visual and motion sensations in a VR simulator.

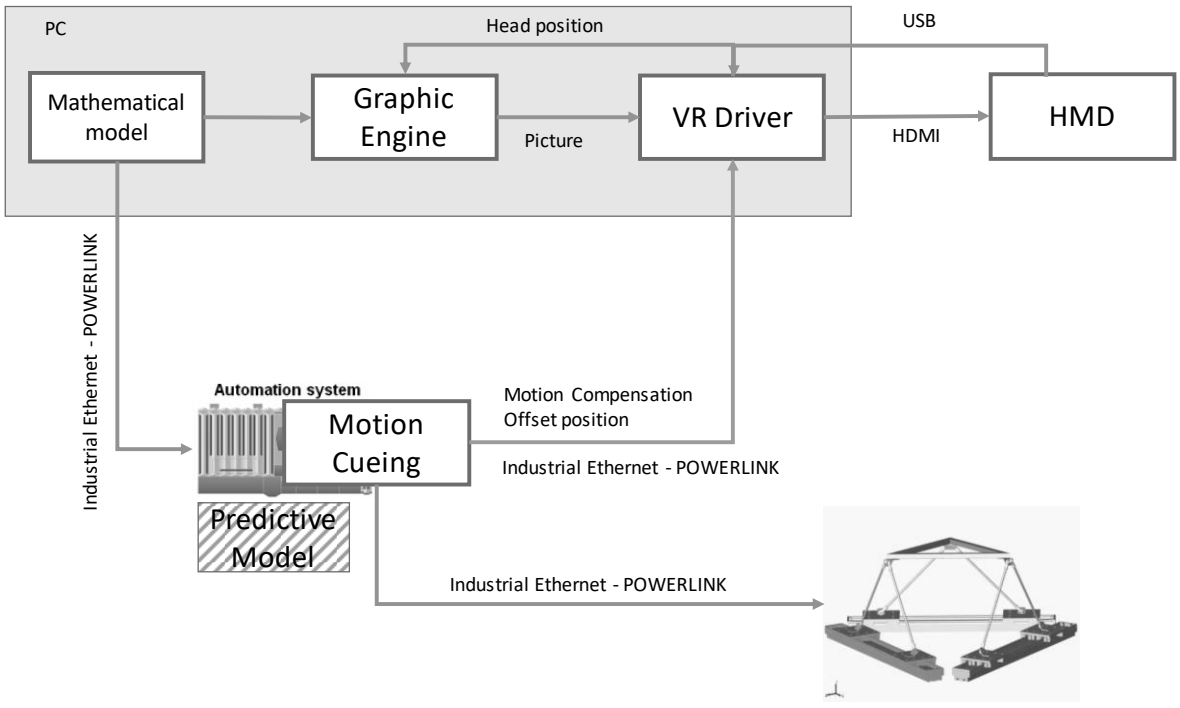


Fig. 5: Motion Compensation.

The head position with motion compensation can be expressed using a transformation matrix:

$$T_{head} = \begin{bmatrix} R_{HMD}R_{offset} & \Delta_{HMD\_0} + R_{offset}(\Delta_{HMD} - \Delta_{HMD\_0}) + \Delta_{offset} \\ 0 & 1 \end{bmatrix},$$

where R_{HMD} and Δ_{HMD} are rotation and translation respectively measured with sensors in VR headset, Δ_{HMD_0} is the origin of the coordinate system to which the measurements relate,

$$R_{offset} = R^{-1}_{motion\_platform}$$

is the inverse rotary matrix of the motion platform,

$$\Delta_{offset} = (I - R_{motion\_platform})h$$

is the inverse translation of the motion platform and h is the head position relative to the center of rotation of the motion platform.

PREDICTIVE MODEL

The prediction model aims to create a time sequence of reproduction of individual perceptions, as shown in Fig 4. These are mainly motion perceptions whose reproduction must ideally be placed before visual perceptions, or at least minimize the time delay.

The prediction model can be solved at various levels of the simulation structure:

1. Mathematical model of a simulated vehicle
2. Motion cueing / washout algorithms
3. Kinematic (position) control of a motion platform

In case 1, additional information such as terrain database data or status of steering components can be used for prediction. However, the computational complexity of this prediction in the order of tens of ms is very high.

The model-based predictive motion cueing algorithm (2) was successfully implemented and presented in the publication (Thöndel 2013). The basis is the implementation of an algorithm in the form of a linear washout filter (MCA) and the search for an optimal solution at a given prediction horizon (see Figure 6). The MPC algorithm predicts at each step the behavior of a controlled system based on its mathematical model. This predicted development is used to detect potential conflicts with the motion platform's kinematic limits. In this case, constraints for maximum deflection in a given direction are updated. At the same time, it is possible to use the predicted development for the time shift of reproduction of motion effects.

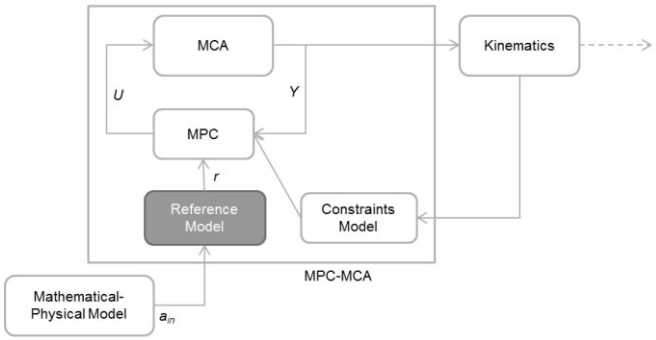


Fig. 6: MPC-MPA Algorithm.

The advantage of this solution is the inclusion of kinematic limits of the motion platform.

The reference model in Fig. 6 represents a prediction of the behavior of a quantity obtained from a mathematical model of a simulated vehicle. In the implementation (Thöndel 2013), this model was simplified to a constant course at the prediction horizon. This simple model was sufficient to calculate the prediction of conflicts with the kinematic limit of the motion platform but is insufficient for the predicted reproduction of motion sensations. This model has been extended using optimal extrapolation based on the recursive least squares method.
The basis of this method is finding the optimal quadratic interpolation function on known historical data and subsequent extrapolation at the prediction horizon.

This prediction model for the prediction horizon p can be expressed by a linear matrix function:

$$r(k) = \underbrace{\left[T_{k=2p}(A^TA)^{-1}A^T\right]}_{const}[a_{in}(k-p) \quad \dots \quad a_{in}(k)]^T$$

where, $(A^TA)^{-1}$ is covariance matrix, T_k is vector $[k^2 \quad k \quad 1]$ and matrix $A = \left[T_0^T \quad T_1^T \quad \dots \quad T_{p-1}^T\right]^T$.

Optimal kinematic control of a motion platform based on a prediction model has been successfully implemented and presented by the author (Thöndel 2017). The main goal was to find an optimal control strategy with respect to energy regeneration and utilization of stored energy in case of power failure. The algorithm uses a linear description of the electromechanical actuator and, at a given prediction horizon, seeks optimal control with respect to constraints. Mathematically, the problem is formulated as minimizing a loss function, which respects the control deviation, the control voltage magnitude, and the power consumption:

$$J = \boxed{(ref - Y)^T Q (ref - Y)} + \boxed{\Delta U^T R \Delta U} + \boxed{U^T S I_i}$$

Error between the predicted value Y and the set-point signal ref — Control voltage — Actual power

The series connection of these two linear models allows prediction and optimal control of a motion platform while respecting the kinematic and dynamic limits. An example of the prediction at the 150 ms horizon is shown in Fig. 7. Data

were obtained by measurement on a developed VR flight simulator.

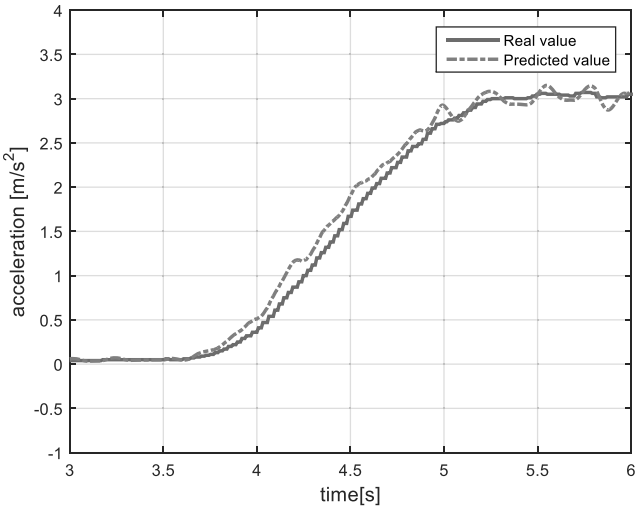


Fig. 7: Simulation result

Computational demands with respect to the required prediction horizon length (150 ms) remain high. As part of the implementation, we are looking for further simplification so that the whole algorithm can be solved in real-time at a high refresh rate (250 Hz).

IMPLEMENTATION

The described algorithm was implemented and tested in a simulator with a motion platform with two degrees of freedom (Figure 8). The test results confirm a high degree of simulation fidelity. The extension of the simulator with the reproduction of motion effects and proper synchronization with other sensations contributed to reducing the risk of the onset of virtual nausea. Proper reproduction of motion effects plays a more critical role in a VR simulator than in a classic simulator and is an essential part of it.

Fig. 8: Prototype of a VR Simulator

CONCLUSION

The article summarises the possibilities of using modern virtual reality in professional simulators with a motion platform. Conditions and potential obstacles for successful implementation are summarised here. The above-described method has been implemented in a simulator with a two-degree-of-freedom motion platform.

The follow-up work focuses on the verification of VR simulator properties with a multi-degree of freedom motion platform. An implementation with a motion platform with six degrees of freedom is planned.

AUTHOR BIOGRAPHY

EVŽEN THÖNDEL was born in Prague, Czech Republic, and studied at the Czech Technical University in Prague, acquiring a Master's degree in Engineering Cybernetics in 2004 and a Doctor degree in Electrotechnology and Materials in 2008. He is a long-time member of the University's teaching and research staff. Besides his academic career, Evžen has been working as a simulation engineer and developer at Pragolet since 2009.

REFERENCES

Ihemedu-Steinke, Q. C., Sirim, D., Erbach, R., Halady, P., & Meixner, G. (2015). "Development and Evaluation of a Virtual Reality Driving Simulator." In Mensch & Computer Workshop band (pp. 491-500).

Meixner, G., & Müller, C. (Eds.). (2017). "Automotive User Interfaces: Creating Interactive Experiences in the Car." Springer.

Kemeny, A. (2014, April). "From driving simulation to virtual reality." In Proceedings of the 2014 Virtual Reality International Conference (p. 32). ACM.

Thöndel E. "Modelling and Optimal Control with Energy Regeneration of a 6DOF Motion Platform with permanent Magnet Linear Actuators". 31st European Simulation and Modelling Conference - ESM'2017, October 25-27, 2017, IST, Lisbon, Portugal

Thöndel, E. "Model Predictive Motion Cueing Algorithm for a Truck Simulator." 27th annual European Simulation and Modelling Conference. Vol. 23. No. 25. 2013.

LOGISTICS SIMULATION

APPLICATION OF MIXED SIMULATION METHODS IN MODELLING PORT-TRAFFIC WITHIN CONGESTED AREAS

Ehiagwina O. Augustine
Maritime and Mechanical Engineering
Liverpool John Moores University
Liverpool, UK
a.o.ehiagwina@2015.ljmu.ac.uk

KEYWORDS

Agent-based, Discrete-event simulation, Port-traffic modelling.

ABSTRACT

This study uses mixed agent-based and discrete event models to simulate maritime traffic in a congested port area. The movement of vessels and tides is simulated using an agent-based model which interacts with port process events generated by discrete event model. This mixed modelling method is implemented using AnyLogic. The agent model is built around objects which can be interfaced to interact with each other by passing messages or continuous time variables. An object may encapsulate other objects to any level of detail. Vessels are modelled as objects and their movement and interactions with other agents is expressed through state charts. This paper focuses on the model construction and initial testing using real world AIS data. The initial results suggest similarities between the real-world data and the model output.

INTRODUCTION

Agent-based microsimulation in which actors within a system are modelled and allowed to interact, are increasingly used in transport research. The behaviour global system emerges from the local behaviour of the individual components and their interactions.

Mixed-model simulation combines models representing individual actors and specific processes. A model of vessel traffic within port might include models:

- Different vessel types each with its own characteristics;
- Interaction between vessels, the environment, and port;
- Environmental models including currents, tides and weather;
- Port processes i.e. availability berths, locks, and anchorages.

Vessel traffic can be modelled by agent-based models which reflect movement characteristics of individual agents and interactions between them such as collision avoidance. These interactions are sometimes deficient in many existing vessel traffic simulation models (Burmeister, Haddadi and Matylis, 1997). Discrete-event simulation models systems as networks of queues and activities, where state changes in the system occur at discrete points of time (Alessandri, Sacone and Siri, 2007; Angeloudis and Bell, 2011; Dragović, Tzannatos and Park, 2017). Port processes are better modelled using discrete event simulation as these follow a sequence of operations performed on agents. AnyLogic is a simulation tool which permits the use of both agent and discrete event simulation. It also has capability of integrating a GIS map into the model. This enables the development of a more realistic and multi-functional model of port-traffic.

This paper explains the development of a port-traffic model using a mixed methodology of agent-based modelling for vessel and discrete-event simulation for modelling of the port. It includes an explanation of object orientated agent based traffic model, the discrete event model used to simulate the port processes, and some initial model output.

AGENT-BASED CONSTRUCT OF SHIP MODEL EXPLAINED USING OBJECT-ORIENTED PROGRAMMING

Agent-based modelling is supported by object-oriented languages such as Java which is used within AnyLogic. Data and programming methods or procedures are encapsulated in objects (agents) that can manipulate their own data and interact with other agents, (Schank, 2013).

In modelling ship movement within the port the starting point would be to define a class of object called “SHIP”. Associated with this object are rules which define how a SHIP interacts with

other agents and the environment. Vessels differ in size, type, propulsion, speed, manoeuvring characteristics etc. The mechanism of inheritance in object-oriented languages simplifies the creation of populations of agents with different characteristics. Instead of writing a completely new class for each type of vessel the structure and properties of the class “SHIP” can be passed to new objects and a new vessel type created by modifying or adding properties.

Encapsulated within the class of objects SHIP is a navigational model which controls vessel movement and its interaction with other agents. This model is implemented as state machine and contains the rules governing the SHIP interaction with other agents. These are based on Collision Regulations (IMO COLREGS) and the manoeuvring characteristics based on the vessel’s physical handling characteristics. The real world movement of a vessel within the port is acquired from AIS data plotted over a GIS map. This allows visualizing of agent movement to identify patterns, and trends.

In modelling the port features, each aspect of the port process (i.e. berth, locks, anchorage, etc.) is defined as objects encapsulated with discrete processes defining that particular port feature.

MODEL CREATION OVERVIEW

Creating port-traffic model includes modelling: (1) the active agents (vessels, port, and environment); (2) agent behaviour (navigational model); (3) the spatial area in which the agents interact using GIS mapping.

Event and time-driven behaviour of an agent is controlled using a state-chart. This is a formal and logical representation of states and transitions that may occur during the dynamic performance of the system under consideration. A state is the condition of an object in which it performs some activity or waits for an event. The state chart represents different contexts in which system behaviours occur. A transition denotes a switch from one state to another. Transitions are relationships between states, drawn as arrows, optionally labelled by a trigger that causes actions. Transitions may be triggered by user-defined conditions (timeouts or rates, messages received by the state-chart, and Boolean conditions).

In this model a behavioural state-chart was created from states linked by transitions used to characterise the different status of an agent and their relationship with other agents. The SHIP behaviours (movement pattern, collision avoidance, interactions with the port etc.) are controlled using the behavioural state-charts, while the port agents representing specific locations like port terminals, anchorage areas, etc. are controlled using a discrete event simulation (Figure 1).

The state chart generates the activity plan of an agents. Each agent has its own passage plan to achieve its destination. Although, **c**onfined to channels within the port, vessels have similar passage plans but with unique trajectories i.e. different waypoints. Most port channels have traffic separation schemes which differentiate inbound and outbound vessel lanes. Inbound and outbound are specific expression for the direction of vessel passages. Historic AIS data of vessel trajectories were plotted to visualise vessel movement pattern and lane positions for both inbound and outbound vessel. Using the width of each vessel lane, GIS-Target lines were created for specific positions identified within the marine map of the port channel. These target lines were used to create target areas which provide clusters of waypoints along the channel between which SHIP agents move. Traffic data for example speed, generated by agents moving through each area is collected and compared with historic data for similar vessels in the same area.

AnyLogic provides a GIS space which serves as a visual environment. This was constructed using spatial data from digital charts of the marine area. The environment is fitted with effects such as tides modelled as agents using state-charts. The comprehensive behaviour emerges as a result of interactions of distinct individual behaviours.

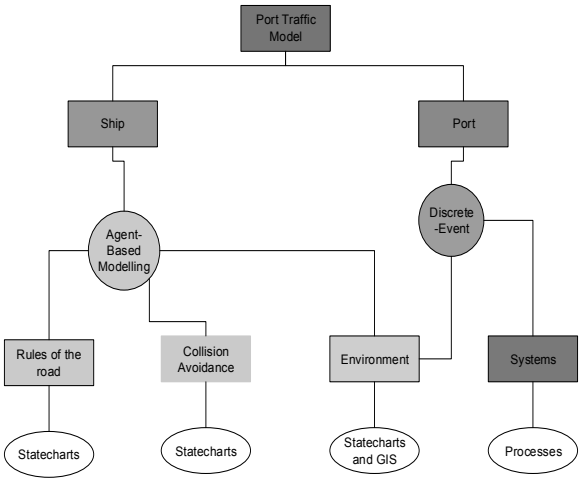


Figure 1: Port traffic model building blocks

MODEL OPERATIONAL OVERVIEW

Figure 2 shows the model operational overview, which includes, the agents, geographical space and their interactions.

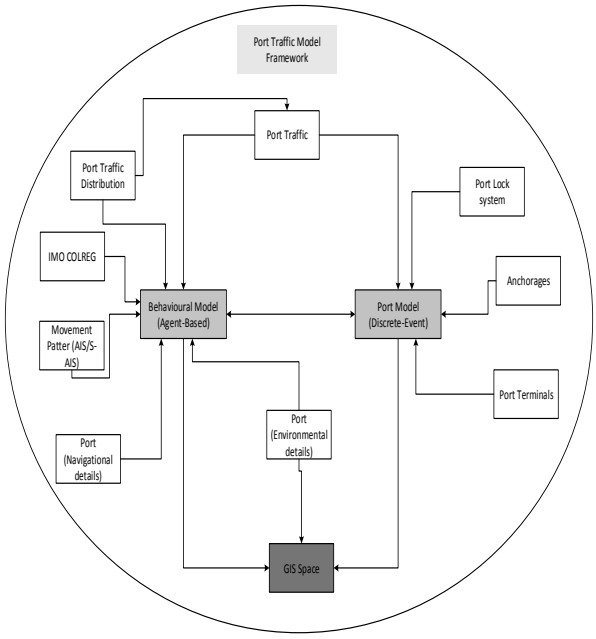


Figure 2: The Model operational overview

SHIP agents are injected into the model based on a distribution of vessel types and respective arrival times created from historic AIS data. The SHIP model is placed within a geographical area based on a GIS map. The agent then interacts with the ENVIRONMENT agent which models weather and tides.

The SHIP, ENVIRONMENT agents interact with the PORT entities which generate events from discrete event simulation models. When a vessel arrives at a port, it immediately informs the port of its arrival, and its next action is based on the port response (if there is an available berth), that is, whether the ship goes for anchorage or proceeds to the berth available. When a vessel is moving within the port, it must oblige to the port, and IMO rules of the road. This movement and interaction with other vessel is controlled by the agent-behavioural and the collision avoidance state-chart within the model. The behavioural state-chart constitute of several states linked to a particular GIS-target-line. The state-chart was developed using the movement pattern observed from the analyse AIS data via trajectory plot, interpreted using "a stop and move" approach proposed by (Spaccapietra et al., 2008), which was further interpreted using a descriptive method. Each state contains behavioural rules and specific ship characteristics governing the vessel movement pattern acquired from the historical AIS data.

Design of Vessel Behaviour

Emergence: As vessels manoeuvre through the port from one target area to the other, they produce emergent behaviours as they draw closer to different areas within the port, which result in speed changes throughout the waterway. For example, as vessels proceed to the Anchorage area, they decelerate, till they come to a stop.

Adaptation: A simple adaptive behaviour built in the simulation is that the ship can always adapt its heading according to the geographical shape of the waterway. Since the model was built on a GIS space, shipping routes were not created, thus; ships are expected to adapt to the directional changes as they proceed to their next target area within the GIS space.

Sensing: The vessels are able to sense their environment along the waterway for other ships and specific areas using embedded rules. For example, vessels are able to observe the tidal conditions, and they immediately proceed to an anchorage point on arrival to the port once the tide is below average high water mark.

***Interactions*:** There are three kinds of interactions. One is the vessels behaviour between one target area to another, creating their own paths, and changing their speed as they manoeuvre through the port. Second is the vessel to port interaction, in relation to tidal conditions, available berth and locks facilities. Third is vessel to vessel collision avoidance interaction.

The agent collision avoidance model was created using a state chart. Each state chart consists of an action a vessel takes depending on what operational mode they are. Actions include changing speed at specific locations such as stopping at anchorage, changes of course and speed during collision avoidance. An agent can get its position, calculate its speed, identify other vessels and their position and speed, the environmental condition and its next position along it planned route (waypoint). From this information the agent makes its decision to move from one point to another and at what speed depending on the agent's state and if environment stays the same. However, in cases of tidal effect, each agent responds to the environmental changes by obliging to the port rule generated by the port model requesting ships to remain at anchorage.

MODEL APPLICATION IN PORT OF LIVERPOOL

The port of Liverpool was used for this simulation study. Measures of vessel traffic activity within the port were generated using historic AIS data and used to validated the global model. The measures include, speed profile, traffic flow, and concentration (density). Flow rate, refers to the number of vessels counted, divided by the elapsed time. Flow rates are collected directly through the entry point, and are measured over time. Traffic concentration refers to the number of vessels along a distance. In this study, it includes both density and occupancy. The first is a measure of concentration over space; the second measure concentration over time of the same vessel stream. Speed is calculated as the average distance over time over the target area. The result of five vessels from the historical AIS dataset in Figure 3 show speed profiles generated by vessels visiting the port without anchoring. Figure 4 shows speed profile generated by the agent based model for similar vessels visiting the port without anchoring, compared with the speed profiles generated from real world data. Figure 5 shows a similar comparison but for vessels which reach the berths from anchorage.

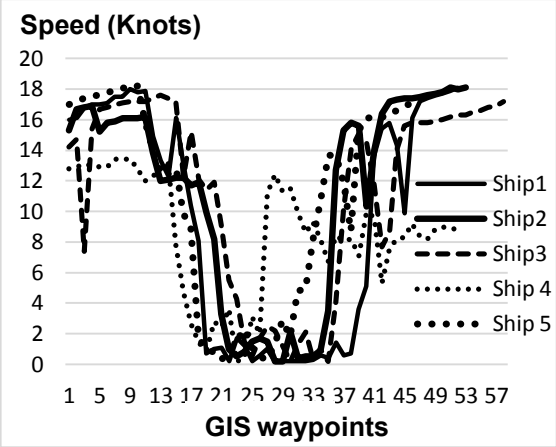


Figure 3: Real-world speed profile from five vessels

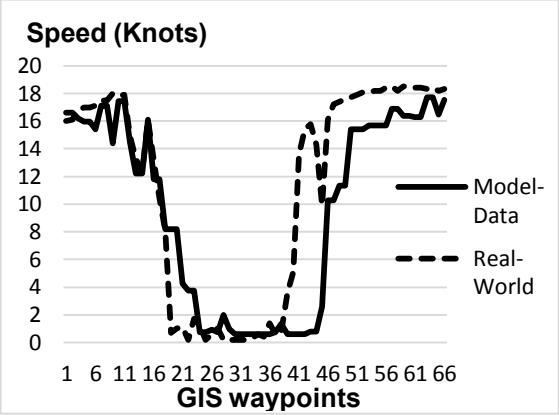


Figure 4: Model speed profiles compared with speed profiles from real-world data across target regions (Without Anchorage)

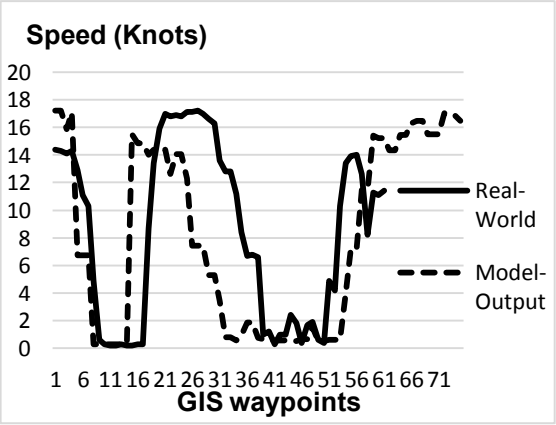


Figure 5: Model-Output Vs Real-World Data to Identify Similar Speed Profile-Pattern (With Anchorage)

CONCLUSION

This paper presents the development of the port-traffic created using a mixed-modelling approach consisting of an agent-based and a discrete-

event models within a GIS environment. The model is at a very early stage and requires further work to develop some of the key features such as elements of collision avoidance model to improve the model fidelity. An agent-based approach is flexible and easily extendible to greater detail. Micro simulation models have the potential to capture the rich and diverse behaviour of agents. The behavioural state chart of agents helps represent complex behaviour and support integration of high-level interaction between ships, ships and port, and with the environment.
The emphasis of this paper is on the proposed methodology for simulating and investigating port-traffic in a congested area. The results presented are at a very early stage. However, they demonstrate the potential modelling methodology.

FURTHER WORK

Future work includes: further enhancement of the collision avoidance model to better model vessel behaviour; more extensive validation and testing across a range of traffic flow measures; the modelling of the different vessel types; and enhancement of the SHIP agent model to include for example propulsion systems which could aid in investigating environmental impacts like emissions.

REFERENCES

Alessandri, A., Sacone, S. and Siri, S. (2007) Modelling and Optimal Receding-horizon Control of Maritime Container Terminals. *Journal of Mathematical Modelling and Algorithms*, 6 (1), 109-133.

Angeloudis, P. and Bell, M.G.H. (2011) A review of container terminal simulation models. *Maritime Policy & Management*, 38 (5), 523-540.

Burmeister, B., Haddadi, A. and Matylis, G. (1997) Application of multi-agent systems in traffic and transportation. *IEE Proceedings-Software*, 144 (1), 51-60.

Dragović, B., Tzannatos, E. and Park, N.K. (2017) Simulation modelling in ports and container terminals: literature overview and analysis by research field, application area and tool. *Flexible Services and Manufacturing Journal*, 29 (1), 4-34.

Schank, J. (2013) Agent-Based Modeling. Methodology and Philosophy. Publisher: WorldPress.org.

Spaccapietra, S., Parent, C., Damiani, M.L., de Macedo, J.A., Porto, F. and Vangenot, C. (2008) A conceptual view on trajectories. *Data & knowledge engineering*, 65 (1), 126-146.

DYNAMIC SCHEDULE EXECUTION TO IMPROVE ADULT EMERGENCY DEPARTMENT PERFORMANCE IN REAL-TIME

Sarah Ben Othman, Dr.
Slim Hammadi, Pr.
Ecole Centrale de Lille, France

Hayfa Zgaya, Dr.
Jean-Marie Renard, Dr. PH.
Université de Lille, Lille, France

Mariagrazia Dotoli, Pr.
Politecnico di Bari, Bari, Italy

KEYWORDS
Emergency Department, Scheduling, Overcrowding, Response time.

ABSTRACT

An Emergency Department (ED) is a very complex system involving heterogeneous patients and several kinds of resources that evolve within a sophisticated process. The management methodology should be chosen in a more effective and targeted way so as to meet the increasing patients' requirements. Our objective is to find out fast solutions for unscheduled arrivals, dynamic competing priorities and heterogeneous patient care needs. The primary objective of this article is to provide ED managers with internal cost-effective solutions and perceptions in order to reduce overcrowding phenomenon impacts and enhance ED performance. Simulation results show that our scheduling method can significantly reduce the total response time of patients.

INTRODUCTION

The Emergency Department (ED) is the heart of hospitals. It provides time-critical care to scheduled and unscheduled patient arrivals. The number of patients arriving with various pathologies at EDs is random and is continuously increasing and the EDs are often incapable of meeting patients' needs (Hoot and Aronsky, 2008) (Niska et al., 2010) (Harrison and Ferguson, 2011) (Abo-Hamad and Arisha, 2013). Overcrowding in EDs remains a significant problem in many countries around the world. It causes serious problems for Medical Staff (MS) members and patients. In this situation, performance indicators become worse such as excessive Waiting Times (WT), increased ambulance diversion, increased total Length Of Stay (LOS) which is the time between patient arrival and his departure, increased medical errors, increased rate of patients who leave without being seen by a doctor (Salway 2017). Care quality is hard to predict in the ED because the ED environment is stochastic; the number of arrivals cannot be predicted, and patient's health state worsening can occur at any time. In addition, the three most interactive actors (patients, MS, and managers) have different outlooks on health care service quality (Cameron et al. 2011). In the literature we can find many publications on modeling and analysis for health care organizations in general and ED in particular. The most used quantitative methods for health care go for mathematical modeling, statistical analysis and simulation (Brailsford et al., 2009). Recently, many studies have been focused on health care systems scheduling. Previous research in hospital logistics can be categorized into two main groups: patient scheduling and resource (e.g. MS) assignment. Patient scheduling is known to have a main role on the performance of the health care system by reducing the WT of patients and MS workloads (Luscombe and Kozan, 2016). A hybrid tabu search algorithm for reducing the total penalty incurred for assigning patients to a non-adequate care box was proposed by Demeester et al. (Demeester et al. 2010). Min and Yih have studied a dynamic programming approach for patients' scheduling in operating rooms considering patients' priority (Min and Yih, 2010). Azadeh et al. have focused on scheduling prioritized patients in EDs using a mixed integer linear programming model and made comparison with Genetic Algorithm (GA) (Azadeh et al., 2014). They proposed a linear structure of chromosomes based on an ordered sequence of operations. This motivates our attention on the effectiveness and efficiency of MS management within the ED. The main objective of ED is to assign the adequate resources to treat the unscheduled and already scheduled patients and calculate dynamic priorities while considering heterogeneous patient care needs and the evolution of their pathologies. The aim of this paper is to propose a resource allocation and health care task scheduling model that represents the ED environment. The scheduling problem in ED can be divided into two interconnected sub-problems: assignment problem and sequencing problem. The MS assignment is modeled as a parallel machine assignment environment with machine groups representing the MS teams. Care tasks scheduling is represented as a flexible job shop problem which is NP-hard optimization problem. The flexible job shop is defined by a set of jobs and a set of resources where each job has a set of operations under precedence constraints. Each task is performed by a single resource selected from a set of resources. In the ED each patient is represented by a job with precedence constrained treatment tasks. The resources are divided into different groups: human resources and material resources.

The objective of this work is to jointly assign one MS member or a medical team with different skills and

schedule the health care tasks while minimizing the response time R_t, the LOS and WT of patients, and the occupancy rate of MS members.

METHODS

Notation

In this section we present the notation required to solve the problem of patients' scheduling in ED.

Sets:

MS denotes the set of MS members in the ED

J denotes the set of all patients

Indexes:

m The index refers to a MS member**.**

t The index of different medical teams t=*1,…,T***.**

A medical staff member $m \in MS_t$.

j The index of an individual patient.

Parameters:

a_j Arrival time for patient $j \in J$. Arrival and triage occur at the same time.

w_j Priority weight for patient triage.

$o_{j,n,t}$ Denotes a health care operation where *j* is the patient, *n* the index of the medical task for that patient, *t* the MS team required to execute the task.

$\{o_{j,,,}\}$ is the set of all the treatment tasks to be executed on patient *j*.

n_j denotes the total number of treatments tasks for patient *j*.

$e_{j,n,t}$ is the execution time of the task $o_{j,n,t}$ by the MS member who has the adequate skills for the treatment of patient *j*.

Decision variables:

X_j denotes the time that the MS member having the needed skills is assigned for patient *j*. At time X_j the treatment tasks $\{o_{j,,,}\}$ become known.

Y_{jn} denotes the starting time for the health care task $o_{j,n,t}$

An optimized assignment and scheduling solution to the ED is to assign each MS member who has the adequate skills to treat patients and schedule all the treatment tasks for all patients. The first objective is to minimize the response time between patient arrival and MS assignment and to minimize the total care task time of all patients. The response time weights w_j are calculated according to the ED MS workload.

Response time $R_t = \sum_{j \in J} w_j (X_j - a_j) + C_{\max}$ where $C_{\max} = \sum_{j \in J} (c_j - X_j)$ with c_j is the completion time of patient *j* and is calculated as follows: $c_j = \max_{n \in \{o_{j,,,}\}} \{Y_{jn} + e_{jnt}\}$

After all the treatment tasks are accomplished for patient *j* $\in J$, by the MS member $m \in MS$, this latter is released and becomes available to be assigned again to a new patient even if he belongs to an already assigned medical team.

The waiting time for each patient is $WT_j = c_j - a_j$ and his length of stay is calculated in real time as follows $LoS_j = t - a_j$.

$$OF = Min(R_t + \overline{WT} + \overline{LoS})$$

Assumption: Minimizing *Rt, WT* and *LoS* leads to smooth peaks of activity in the ED and reduce overcrowding situations.

Proof : According to meetings with MS of Adult Emergency Department of the Regional University Hospital (RUH) of Lille and many visits to emergencies, we can define an overcrowding situation in the ED as an imbalance between the care load and the capacity of patients' treatment over a sufficient length of time that may lead to adverse consequences for proper functioning. We analyzed a sample of data provided by the ED of Lille RUH. We identified indicators of load, capacity and consequences. The supply and demand measurement and analysis show that there is a close connection between the overcrowding situation and the indicators sited above.

SCHEDULING APPROACH

The algorithm is designed in one hand to optimize the assignment of MS to patients considering the needed skills and to obtain a good solution for patients' scheduling minimizing the objective function OF on the other hand. As already mentioned, the ED scheduling can be considered as a flexible job shop problem, real-sized problems implicate complex and large search space and cannot be solved by exact methods within a reasonable amount of time. GAs have been demonstrated to be efficient for the exploration of complex research space. GA belongs to the family of evolutionary algorithms

inspired by natural evolution the principle of "the fittest can survive". Inspired by molecular biology, GA considers a solution as a chromosome structure containing good and bad phenotypes. Assuming that the right phenotypes are part of the optimal solutions, GA uses the mechanism of reproduction, natural selection and the mutation principle to produce new solutions in which the chromosome contains better phenotypes.

Algorithm 1: the GA approach

Output: a set of N good scheduling solutions S^*
Begin
Generate a set for random solutions to create the initial population
while (stop criteria are not reached) **do**

- Calculate the fitness function value OF
- for each chromosome
- Select 2 parents *P1* and *P2* at random
- Apply a controlled crossover algorithm with a probability p_c, in order to a predetermined percent of individuals selected from the population
- Apply a controlled mutation algorithm with a probability p_m in order to a predetermined percent of individuals selected from the population
- Select N new individuals (survivor selection) and build a new population S^*
- Update the stopping criteria

end while

for each un-scheduled task $o_{j,n,t}$ do

Assign_and_schedule ($o_{j,n,t}$, S^*)

end for

End

Chromosome representation

The most important steps in designing a GA is the manner by which the structure of the chromosomes is represented and then inferred. In the proposed algorithm, a cubic chromosome representation is used which encodes a schedule as a three-dimensional space. Three axes are used to generate a chromosome: "MS", "patients" and "time". The time axis is divided into intervals with different sizes. The scheduling horizon is discretized into many periods with different duration. Each period contains many slots with the same or different duration.

Crossover operator

The crossover is a critical operator of GAs which leads to produce offspring by doing effective combination of chromosomes. The controlled crossover schema is used in the proposed GA in order to move the starting time of the patient treatment process forward or backward for a given MS member. Only the "time" and "patients" axes are considered.

Mutation operator

Mutation allows us to create subtle changes in the chromosome in order to increase the diversity of the current generation of population to the next one. In the proposed GA we use a controlled mutation for example changing in one slot the assignment of MS member by controlling his availability (handling the "MS" axes) or to switch the two periods of two patients if it is possible (handling the "patients" and the "time" axes).

Algorithm 2: Assign_and_schedule

Output: Assign the task $o_{j,n,t}$ to an available resource *t* who has the adequate skills to treat the care task *n* of patient *j*
Begin
swich(t)
case (doctor)
select $m \in MS_t$ with lowest load which has a needed skill
set *m* as physician to treat the patient *j*
Insert $o_{j,n,t}$ in sequence S^* respecting the priority FIFO rule
case (medical team)
select all $m \in MS_t$ with needed skills to compose the team *t*
set t as medical team to treat the patient *j*
Insert $o_{j,n,t}$ in sequence S^* at position *q* respecting the resource constraints such as the availability of all the medical member of the team *t*
case (default: X-ray, ….)
Select a needed resource with lowest utilization
Insert $o_{j,n,t}$ in sequence S^* respecting the priority FIFO rule
End_swich
End

Dynamic environment

The optimal solution generated by the proposed GA is designed to be monitored in a dynamic ED environment where patients arrive randomly without prior order and have to be included in the current schedule. When an event occurs in the ED which is either a new patient arrival or a doctor available (resulting from the end of patient treatment) as shown in the algorithm 2 (Assign_and_schedule), a reschedule of the task-resource sequences is necessary. Before rescheduling occurs, we have to check the availabilities of MS and update the current set of treatment tasks so that it contains the

remaining tasks and the new tasks related to new arrivals. All commenced and completed tasks are removed from the set of tasks to be rescheduled.

SIMULATION AND RESULTS

Our experiments consist in a case study involving Lille Regional University Hospital (RUH) Adult Emergency Department (AED) in the north of France that receives an average of 9400 visits a year. It has 10 beds in short stay hospitalization unit, 8 consultation boxes in outpatients' unit, a suture room, a plaster room, an emergency room and 2 waiting rooms. Upon admission to the AED, each patient always receives a first evaluation of his health state by a nurse at the reception. Patients' triage according to their emergency degree makes it possible to set them on the adapted care pathway: The Short Circuit (SC) or the Long Circuit (LC). SC is characterized by a pathway and dedicated staff. SC allows speeding patient whose health condition is considered less severe through the process of care delivery. Complex situations requiring complementary exams (biology exams, RMI, etc.) are directed towards the LC. Finally, patients whose health condition is considered precarious are directed, upon arrival, to a vital emergency room for immediate care.

Real-size scenario

To evaluate the proposed algorithm for deployment in the AED we develop a real size scenario. At his arrival, a priority degree and a pathway are assigned to each patient based on the real functioning of the AED of RUH of Lille. We consider that patients arrive according to a Poisson process. The arrival rates and the resources' counts are based on the data from the AED case study. We have an extract of the data concerning all the patients passed through the emergencies, their pathologies, their arrival time as well as their exit time and the exams that were assigned to them over a year. From this database we can determine some characteristics related to the patient journey in the AED such as the average annual number of visits that reaches 9610 visits as well as the average patient waiting time (2 hours and 30 minutes). MS including doctors and nurses are rostered at unpredictable levels during the day in accordance with expected workload. For the sake of this analysis we have assumed a constant staff level. In the computational study, we have 15 practitioners. Each practitioner can have one or many skills needed for the execution of health care operations.

The inputs of the scheduling algorithm are the lists of patients and corresponding health care operations. The output is a Gantt chart (Figure 1). Each patient is therefore represented by a list of care operations to be performed by MS team. Our scheduling approach is based on 2 phases. The first phase consists of treating the already scheduled patients who arrive at the AED. A GA is applied in this phase. The scheduled patients correspond to the constraints of the GA. The second phase considers the dynamic environment of emergencies which mainly corresponds to the arrival of non-scheduled patients at the AED. The dynamic scheduling algorithm considers patients' pathologies and their emergency degrees. For a given scheduling, the proposed scheduling approach gives the best performing Gantt chart possible, respecting the priority of the patients. The objective is to minimize the time needed to treat all patients (response time). In the computational study there are 100 generations of a population of 60 individuals.

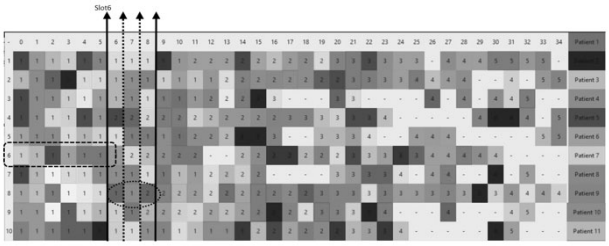


Figure 1. Dynamic schedule generated for MS assignments and resource sequences using our scheduling approach

The populations are composed of the 5 best previous individuals, 25 recombinations of the best individuals and 30 random individuals. In output, we have the best scheduling, its Gantt and its execution time. Figure 1 represents a screenshot of the Gantt chart derived from the simulation of our scheduling algorithm. In this simulation, we consider slots of 5 min and the MS is represented by only physicians. In Figure 1, X-axis represents the time divided into slots and the Y-axis represents the MS members charged of executing the different health care operations corresponding to scheduled, non-scheduled and non-scheduled and urgent patients arriving to the AED of RUH of Lille. Each MS member is multi-skilled (5 skills). Each color in the Gantt chart corresponds to a patient being treated by a MS with the needed skill. For example, at his arrival, patient 7 is treated by the MS 6 using skill 1 during the first slot. In the following slot, MS 6 (skill 1) treats patient 11, then patient 5 then patient 6. In the slots n° 4 and n° 5, MS 6 treats patient 9. The arrival of new patients and a sudden worsening of the health condition of patients trigger a reschedule of the task-resource sequences. Before rescheduling occurs, the list of care tasks to be scheduled is updated. If a health care process corresponding to a patient is completed, the list of available skills must be updated as well. The MS assignments to patients are made according to the weighted response time objective. At time τ (slot 2) a new patient (patient 1) arrives at the AED but there is no

available MS member with the needed skill to treat him, so patient 1 waits until a MS member becomes available, and his treatment begins at the beginning of the slot n°7. Patient 1 corresponds to the highest weighted response time among the set of patients waiting for treatment at time τ. When slot 7 starts, a new patient (patient 10) arrives to the AED with a high emergency degree and must be incorporated into the resource sequences. The Gantt chart has been updated by interrupting the health care process of patient 4 being executed by MS 8 in order to schedule patient 10. In fact, in emergencies, because saving the life of each patient is of vital importance, the interruption of non-urgent care tasks is allowed in order to meet the needs of the more urgent patients. Slot n°7 corresponds to the earliest feasible position in the sequence.

Figure 2. DSS interface showing the evolution of performance indicators in the AED

After the validation of our proposed scheduling approach in collaboration with the Lille RUH AED managers, we developed a Decision Support System (DSS) presenting an ergonomic interface using the language JAVA under Eclipse (Figure 2). In the AED of Lille RUH, there is no DSS implemented in order to better manage the health care process. There is only “ResUrgences” which is a software platform Net Intra/Extranet communicating to manage reception and implementation of operational emergency patient service. For the implementation of our approach in the AED of Lille RUH and the validation of the evaluation and scheduling results, we started by connecting our DSS to “ResUrgences” and we did real simulation scenarios.
Figure 3 shows an extract of a comparison between patient treatment time given by “ResUrgences” (orange curve) only, and results given by “ResUrgences” (blue curve) connected to our approach. We can notice that our scheduling approach has led to the reduction of the treatment time. As a result of this simulation and after many interviews and meetings with the emergency service managers and Berger-Levrault, the publisher of “ResUrgences” software, we found out that the integration of our scheduling approach in “ResUrgences” could be very successful.

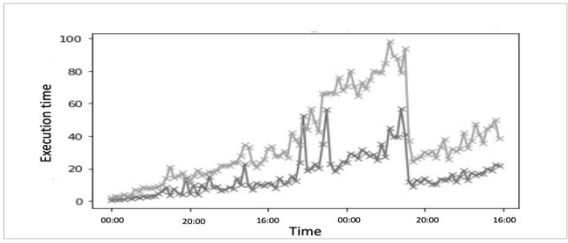


Figure 3. Comparison between results given by ResUrgences and "ResUrgences"+Our Approach

Berger-Levrault intends to proceed with a research and innovation strategy in order to integrate a scheduling module into “ResUrgences” software in order to more simplify the management of emergency services and to increase patients' satisfaction.

CONCLUSION

In this paper, we have proposed a scheduled approach based on two phases: a GA and a dynamic scheduling algorithm. Our method considers the dynamic feature of AED environment and aims to solve the problem of the interferences between scheduled, non-scheduled and non-scheduled and urgent patients. The objective is to minimize the time for treatment process. In our scheduling we take into account the information on the evolution of the state of health of the patients during the scheduling. This information can be given by Smart Things (such as smart watches, smart sensors, etc.) able to give real-time information on blood pressure for example, etc. Because in emergencies, time does not flow in the same way depending on whether you are caregiver or patient and the waiting time usually tends to expand, a theoretical scheduling time is given to patients in order to comfort them and give them an idea about how much time they have to wait before the first consultation. Then the real time scheduling algorithm is applied in order to update the assignments given by the GA The information system in the AED of RUH of Lille (ResUrgences) records arrivals, departures, emergency degrees and pathologies of the patient cohort. The exact information about health care operations and sequencing were missing, so we have collaborated with AED physicians to develop a real scenario for the purpose of testing the dynamic scheduling algorithm. For future work, we intend to include this approach in a multi-agent system. Agents will be able to orchestrate the scheduling in real time during the operating phase in order to optimize the health care process. They will therefore have an optimizing behavior trying to enhance the performance indicators during the scheduling execution.

REFERENCES

Abo-Hamad, W. and Arisha, A. (2013). Simulation-based framework to improve patient experience in an emergency department. *European Journal of Operational Research, 224:154–166*.

Azadeh A., Hosseinabadi Farahani M., Torabzadeh S., Baghersad M. (2014). Scheduling prioritized patients in emergency department laboratories, Computer Methods and Program in Biomedicine. 117: 61–70.

Brailsford S. C., Harper P. R., Patel B. & Pitt M. (2009). An analysis of the academic literature on simulation and modelling in health care. *Journal of Simulation, 3 (3): 130–140.*

Cameron P. A., Schull, M. J., & Cooke, M. W. (2011). A framework for measuring quality in the emergency department. *Emergency Medicine Journal, 28 (9): 735–740.*

Demeester P., Souffriau W., Causmaecker P.D., Berghe G.V. (2010). A hybrid tabu search algorithm for automatically assigning patients to beds. *Journal of Artif. Intell. Med. 48: 61–70.*

Harrison, J. and Ferguson, E. (2011). The crisis in united states hospital emergency services. *International journal of health care quality assurance, 24(6):471–483*.

Hoot, N. and Aronsky, A. (2008). Systematic review of emergency department crowding: Causes, effects, and solutions. *Annals of Emergency Medicine, 52(2):126–136*.

Hupert, N., Hollingsworth, E., and Xiong, W. (2007). Is overtriage associated with increased mortality? Insights from a simulation model of mass casualty trauma care. Disaster Med Public Health Prep 1(1 Suppl): S14-24.

Luscombe R. and Kozan E., (2016), Dynamic resource allocation to improve emergency department efficiency in real time. *European Journal of Operational Research. 255: 593–603.*

Min D. and Yih Y.W. (2010). An elective surgery scheduling problem considering patient priority, *Comput. Oper. Res. 37: 1091–1099*.

Niska, R., Bhuiya, F., and Xu, J. (2010). National hospital ambulatory medical care survey: 2007 emergency department summary. National Center for Health Statistics, 26.

Salway RJ, Valenzuela R, Shoenberger JM, Mallon WK and Viccellio. (2017). Emergency Department (ED) Overcrowding: Evidence-Based Answers to Frequently Asked Questions. *Rev. Med. Clin. Condes. 28(2):213-219*.

BIOGRAPHIES

Sarah Ben Othman is a PostDoc within CRIStAL laboratory in the Ecole Centrale de Lille (Ec-Lille) in France. She is a PhD Engineer in automation. She obtained her PhD degree in December 2015 from Université Blaise Pascal in Clermont-Ferrand and a master's degree in supply chain management in November 2013 from EC-Lille in France. Her current work regards the control of flows in hospital emergency departments, assessing and anticipating crowding mainly in pediatric departments. She developed skills in modeling and optimization, supply chains and decision support systems.

Hayfa Zgaya is an Associate Professor of Computer Science in the University of Lille (France). She obtained her PhD degree in July 2007 from the engineering French High School Ecole Centrale de Lille and her master's degree in November 2002, from Nantes University (France). Recently (October 2014), she obtained the highest academic qualification from University Lille 1 (France), allowing her to supervise research. Her main research areas are the optimization, the artificial intelligence and the logistics issues.

Mariagrazia Dotoli (M'96, SM'2012) received the Laurea degree in Electronic Engineering with honours in 1995 and the Ph.D. in Electrical Engineering in 1999, both from Politecnico di Bari (Italy). She was a visiting scholar at the Paris 6 University and at the Technical University of Denmark. Since 2003 she has been expert evaluator of the European Commission, initially under the 6th and 7th Framework Programmes and now in Horizon 2020. She is currently an Associate Professor in Control Systems Engineering of Politecnico di Bari, Italy, which she joined in 1999 as Assistant Professor. She was the Vice Rector for research of Politecnico di Bari for years 2010-2013. Her research interests include modeling, identification, management, control and diagnosis of discrete event systems, manufacturing systems, logistics systems, energy networks. Prof. Dotoli is an Associate Editor of the IEEE TRANS. ON AUTOMATION SCIENCE AND ENGINEERING and of the IEEE ROBOTICS AND AUTOMATION LETTERS. She is a member of the editorial board of the journals: MEDITERRANEAN J. OF MEASUREMENT AND CONTROL, INT. J. OF AUTOMATION AND CONTROL, INT. J. OF SYSTEMS SIGNAL CONTROL AND ENGINEERING APPLICATION, and INT. J. OF DISCRETE EVENT CONTROL SYSTEMS. She is the Workshop and Tutorial Chair of the 2015 IEEE Conference on Automation Science and Engineering. She has been member of the International Program Committee of 40+ international Conferences. She is author of 150+ international printed publications.

Slim Hammadi is a full Professor of modelization, optimization and control of complex systems at the Ecole Centrale de Lille. He is a director of OSL team of CRISTAL Laboratory and leader of OPTIMA research Group of CRISTAL. He is a Senior member of IEEE/SMC and has served as a referee for numerous journals including the IEEE Transactions on SMC. Prof. Hammadi was Co-Organizer of many Symposia (IMS) of the IMACS, IFAC and IEEE SMC Multiconference. He has organized several invited sessions in different SMC, IFAC, IEEE, WAC, conferences where he was session chairman. He was president of the International congress on Logistic and Transport LT2004, MHOSI2005, LT2006, LT2007 and LT2009. He has published 220 scientific papers (40 on international peer reviewed scientific journals, 12 on chapters of international edited books and 168 in peer reviewed proceedings of international conferences). His teaching and research interests focus on the areas of production control, production planning, computer science, discrete and dynamic programming and computer integrated manufacturing. Email: slim.hammadi@ec-lille.fr.

USING DISCRETE-EVENT SIMULATION TO EXPLORE THE IMPACT OF USER BEHAVIOURS ON THE EFFECTIVENESS OF A TERMINAL APPOINTMENT SYSTEM

Mihai Neagoe
Mohammad Sadegh Taskhiri
Paul Turner
ARC Centre for Forest Value
University of Tasmania
Hobart, Australia
Email: Mihai.Neagoe@utas.edu.au

Hans-Henrik Hvolby
Centre for Logistics,
Dep. Of Materials and Production
Aalborg University
Fibigerstræde 16, 2-109
Aalborg, Denmark

KEYWORDS
Discrete-event simulation, transport management, congestion management, information systems, marine terminals

ABSTRACT

This paper analyses driver behaviours through a discrete event simulation model of a bulk cargo marine terminal. Specifically, punctuality and the proportion of appointed arrivals are examined for their impacts on truck flows and turnaround times at the terminal. Using empirical data from an Australian wood chip export marine terminal the simulation provides useful insights into the terminal's contemporary experience with respect to truck congestion. The simulation results indicate that the proportion of appointed arrivals, used as a proxy for the appointment system use, can improve truck turnaround times primarily in the earlier stages of adoption. Greater truck arrival punctuality can also improve the effectiveness of the appointment system particularly as the system's adoption increases. Shifting the focus from enforcing punctuality towards ensuring use of the system, primarily in the early stages of system adoption, appears to deliver greater benefits for both terminal operators and transporters. Based on this work the terminal operator is now moving rapidly to procure and implement an electronic appointment system as part of its initiatives to mitigate truck congestion at the terminal.

INTRODUCTION

One of the most effective methods to manage congestion and communicate with multiple users at marine terminals is through the deployment of digital terminal appointment systems. This stated, there are disagreements over how best to optimise the benefits that may accrue from these systems. Most researchers advocate optimising system parameters such as:
(a) lead time for appointment time selection, appointment window length (Chen et al. 2013);
(b) number of appointments per time-window (Torkjazi et al. 2018);
(c) appointment spacing, and truck servicing rules (Li *et al.* 2018).
These parameters are usually optimized through a range of techniques including linear programming (Chen and Jiang 2016), queuing theory (Guan and Liu 2009) or through simulation (Huynh and Walton 2011). Unfortunately, in most of these approaches user behaviours are perceived as obstacles, barriers or disruptions to the development of optimal system solutions. In response, some researchers advocate the use of penalty systems to enforce user compliance with the appointment system schedules (Ramírez-Nafarrate *et al.* 2017). Alternatively, others argue for the introduction of more complex truck service rules to maintain a high level of system efficiency (Li *et al.* 2018).
While these enforcement and system rules approaches directly impact on moderating users' behaviours in the models produced, in reality they suffer from a series of practical limitations:
(1) Failure to involve users' in decision-making around system's parameters and functionality is short-sighted and may have unintended consequences in practice (Huynh *et al.* 2016). As Ackoff (1978) argued: *"In problems, the solution to which involve the reactions of others, their participation in the problem-solving process is the best protection against unexpected responses* […] *A failure to consult others who have a stake in our decision is often seen as an act of aggression".*
(2) Real-time access to the necessary information required for complex system rules development and implementation is costly and hard to deliver;
(3) These approaches do not contain a basic working knowledge of the underlying causes behind users' behaviours increasing the likelihood that unexpected outcomes such as system work-arounds or misuse occur (Morais and Lord 2006).
This paper argues that the performance of terminals in terms of truck turnaround times and/or equipment efficiency will be affected by the behaviours of the terminal users. It illustrates that user behaviours can be modelled by their arrival punctuality (Li, *et al.* 2018; Ramírez-Nafarrate *et al.* 2017) or the probability that drivers miss appointments, arrive un-appointed (Huynh and Walton 2008).
Using a discrete event simulation model of a bulk cargo marine terminal, arrival punctuality and proportion of planned appointments, were modelled for their impact on truck flows and turnaround times at the terminal. The modelling results have been used to support the involvement of terminal users in the development of an appointment system, to understand their requirements, consequences on the users' behaviours and the consequences of the users' behaviours on the effectiveness of the system.
Based on this work the terminal operator is now moving rapidly to procure and implement a digital appointment

system, embedding user centred-principles, as part of its initiatives to mitigate truck congestion at the terminal.

FIELD SITE AND DATA COLLECTION

The simulation model is based on a wood chip export marine terminal located in Australia. The terminal handles wood chips in bulk delivered by trucks from mills located in a relative close proximity. Wood chips are stockpiled in preparation for vessel arrivals and loaded on board ships using a conveyor belt system. Wood chips are typically used by international pulp and paper producers or as biomass fuel. The terminal handles approximately 52,000 trucks per year which bring in close to 1,600,000 tons of goods.

Trucks arriving at the terminal for wood chip deliveries are first weighed at a weigh-bridge. Operators then drive to the wharf area where they wait for one of the two hydraulic tipping ramps to become available. Once a ramp is available, operators drive on the ramps and the entire rig is tilted at an angle forcing the wood chips out of the trucks' trailer(s) and into a hopper. The hopper is shared by the two hydraulic ramps and is progressively emptied by a conveyor belt system that delivers the wood chips onto the stockpiles. Once trucks are emptied, the operators drive off the ramp to a weigh-bridge where the empty weight of the trucks is measured. The terminal and supply chain setup have been explored in greater detail in previous work (Neagoe et al. 2018).

The terminal's throughput has been progressively increasing in the past 5 years leading to significant truck congestion. The consequences of this situation range from increasing costs and uncertainty for transporters and customers to increasing labour costs and maintenance scheduling for the terminal operator. The terminal operator and users have considered a range of potential options to address truck congestion and expressed interest in a truck appointment system. The research team therefore developed a simulation model to assist in highlighting the influential aspects that may affect the appointment system's effectiveness.

Data Collection

The input data for the simulation model was collected from two sources: the weigh-bridge database and geo-location data from global positioning system (GPS) units mounted on several trucks that used the terminal. The research team also undertook 2 visits at the terminal and the processing mills to understand the production, transportation and unloading processes and have engaged regularly with the terminal's and users' staff to guide the development of the simulation model. Geo-fences were setup around terminal infrastructure and, using the GPS data, each truck's visit duration in every stage of the terminal unloading process could be determined. The weigh-bridge data were used to determine the inter-arrival times between consecutive truck visits. The empirical inter-arrival times and the geo-fence visits durations were analysed with Arena Input Analyzer to determine the best-fit probability distributions.

Approximately 7 months of truck arrival data (consisting of 15,622 weigh-bridge observations) supplied by the terminal operator and 3 months of geo-location data were used (consisting of 6,709 geo-fence entries). The geo-location data were supplied by one of the transport companies. The geo-fence data were cross-referenced with two time-and-motion measurements to ensure the representativeness of the data.

Table 1: Arena Input Analyzer Fitted Distributions

Data	**Unload**	**Weigh-Out**	**IAT**[a]
Offset	5.5	0	-0.5
Distri-bution	LOGN (5.16, 3.97)[b]	N(3.46, 1.68)[c]	G(1.49, 6.97)[d]
SSE[e]	0.002	0.016	0.002

[a] Inter-Arrival Time; [b] Lognormal distribution; [c] Normal distribution; [d] Gamma distribution, [e]Sum of Square Errors

The fitted distributions output by Arena Input Analyzer are summarized in Table 1. These fitted distributions formed the quantitative data input for the discrete-event simulation model. The fitted distributions' average, spread and shape were also compared with the rule-of-thumb figures the terminal staff used on a regular basis. Importantly, both the data collection and analysis processes were discussed with the terminal and users' staff which led to an iterative development of the simulation model.

TERMINAL SIMULATION MODEL

The conceptualization and development of the terminal simulation model were influenced by discussions with terminal and transporters' staff. The preliminary weigh-bridge data analysis performed by the researchers revealed that whilst the arrival frequency average was approximately 10 minutes/truck, 10% of trucks arrive within 1 minute of one-another and 30% of trucks arrive within 3 minutes of one-another. Consequently, an appointment system was considered by the researchers to increase the spacing between truck arrivals and limit waiting times due to overlapping arrival times. However, the idea of an appointment system was not immediately accepted.

Transport operators, in particular, expressed two main concerns with regards to the potential introduction of an appointment system: the impact of system adoption and the types of behaviours incentivised by the system's rules and/or penalties. Transport operators considered the situation where only some of the drivers would make appointments whilst others would continue to arrive unappointed. The concern raised was that operators with appointments would still have to wait behind unappointed drivers and therefore truck turnaround times would not decrease sufficiently to warrant the effort of using the system. The researchers modelled a non-compulsory appointment system to explore whether the transporters concerns were founded.

The second issue raised by transporters was regarding the potential hazards of incentivising drivers to arrive at a particular time for an appointment. Potentially dangerous driver behaviours (e.g. speeding) could be incentivised if drivers would not be serviced by the terminal, or if they were penalized for early or late arrivals. The expression "*die for a deadline*" was used by transporters to describe the potential danger of this situation. Consequently, a first come-first served service rule was modelled by the researchers to evaluate the potential effects of the appointment system without placing penalties or restrictions on drivers.

As a result of the discussions with terminal and transporters' staff, the two-stage simulation model was developed. The model is implemented in the Python programming language. The two-stage approach separates the truck arrival generation from the truck processing. This is required because of the two emerging requirements from stakeholder discussions: the non-compulsory appointment system and the first-come first-served service rule. Table 2 presents the two stages of the simulation's logic model: the truck arrival generator and the truck processing stage.

Table 2: Simulation Logic Model - Stages and Steps

Step	Stage 1 – Truck Arrival Generator
1.1	Random choice of appointed/un-appointed arrival;
1.2A	**If un-appointed arrival:** arrival time = previous truck arrival time + inter-arrival time;
1.2B	**Else appointed arrival:** arrival time = previous app. time + appointment interval + punctuality;
1.3	Sort truck list based on truck arrival times;
1.4	Generate truck object (payload, capacity, product and arrival times);
Step	**Stage 2 – Truck Processing**
2.1	Next truck object from arrival list;
2.2	Weigh-in at Weigh-bridge (1 min);
2.3	Drive to Unloading area (1 min);
2.4A	**If any unloading ramp free and available:** Unload;
2.4B	**Else no unloading ramp free and available:** Accrue waiting time;
2.5	Drive to Weigh-bridge - 2 min;
2.6A	**If weigh-bridge free:** Weigh-out;
2.6B	**Else if weigh-bridge not free:** Accrue waiting time;
2.7	Calculate service and waiting times per truck;

In the first stage, the truck arrivals are generated. The arrival time of the truck is then calculated based on whether or not the respective truck has an appointment. The truck arrival list is then sorted in ascending order of arrival times and fed to a generator function which creates truck objects with payload, capacity, product characteristics. Truck objects generated in the first stage are processed in the second stage.

The inter-arrival times, weigh-out and unloading ramp service times are stochastically drawn using Monte Carlo from the fitted distributions presented in the previous section. The weigh-in in stage is held constant at 1 minute/truck. the drive times from the weigh-bridge and unloading ramp and back are held constant at 1 and 2 minutes respectively.

The simulation model's design has a relatively low complexity which was a purposeful choice of the research team. The model was primarily used as a discussion element with the terminal and users' staff which had a set of diverse knowledge sets, expertise as well as demographic and socio-economic backgrounds. Therefore, a low complexity design aimed to enhance the accessibility to a diverse audience.

The scope of the terminal simulation model extends only to the truck unloading process. Other terminal operations such as product handling and storage or vessel loading could not be included in the model due to the absence of reliable data.

The truck processing stage models closely the terminal unloading process observed by the research team. The simulation model logic, parameters and output have been discussed with terminal staff to enhance accuracy and validity.

Model Validation

The model's input distribution parameters as well as the model's output were validated against the empirical data.

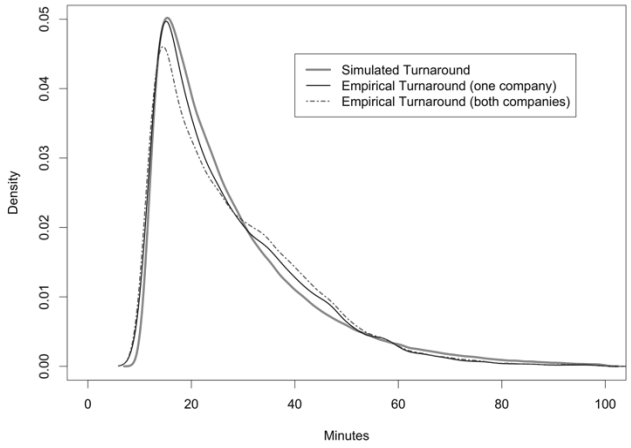


Figure 1: Simulated vs Empirical Truck Turnaround Times

The comparison between the model's output and the empirical truck turnaround times, presented in Figure 1, reveal minimal differences. The solid black line represents the turnaround times of trucks operating for one of the forestry companies, the same that supplied the geo-location data. The blue dashed line represents the turnaround time distribution of both transport companies using the terminal. Data exploration revealed that one of the companies was only recoding the duration of the visit on the weigh-bridge in cases where the visits were over 30-40 minutes, therefore systematically biasing the data. Consequently, for the purposes of validating the model's turnaround time output with empirical figures, data from only one transport company was used.

The average simulated turnaround times are approximately 3.5% lower than the average empirical turnaround times of the transport company taken as a reference point. This difference may be caused by behaviours that cannot be accurately and regularly geo-located and measured such as: driver breaks, differences in drivers' speed and skills or breakdowns in terminal equipment and trucks. These qualitative differences due to human factors could potentially be measured and included. However, the difference between the empirical and simulated figures is very small and acceptable for the model's purpose.

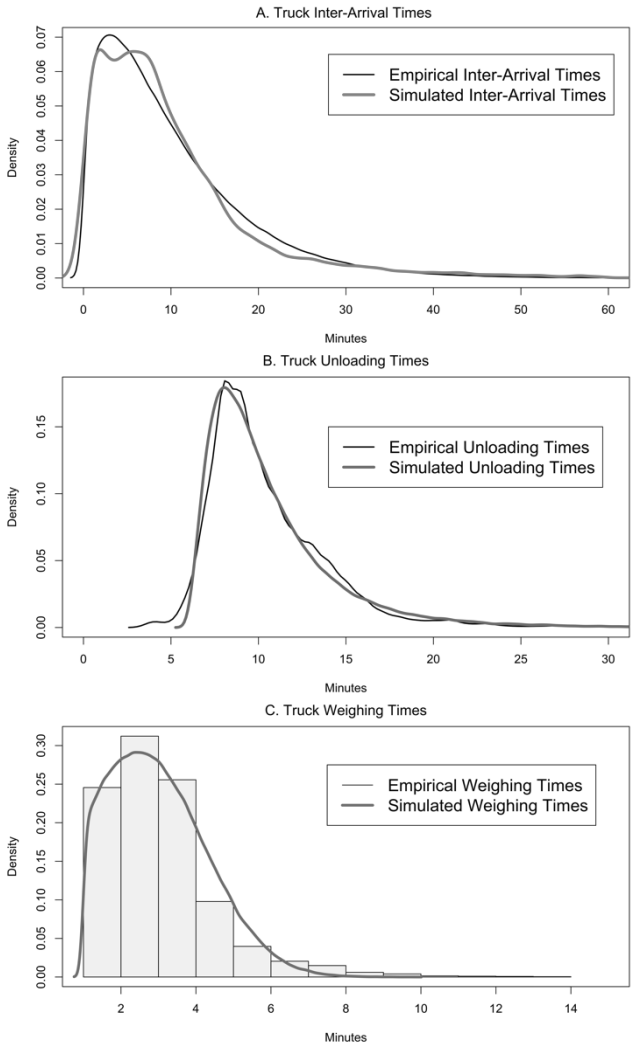


Figure 2: Empirical vs Observed Truck Inter-Arrival Times, Weigh-Bridge and Unloading Times

Figure 2 illustrates the simulated and the empirical distributions. The truck inter-arrival times are shown in Figure 2 A. The empirical inter-arrival times were determined by subtracting subsequent arrivals of trucks in the weigh-bridge records. The simulated and empirical truck unloading times, shown in Figure 2 B, appear to have a larger proportion of low unloading time durations when compared to the empirical distribution. The truck unloading times were determined from the geo-fence entries In Figure 2 C, a similar, although less pronounced pattern, can be seen regarding the truck weighing times.

User Behaviour Scenario Analysis

The driver behaviour factors included in the scenario analysis were punctuality, and proportion of appointed arrivals. These factors were adapted from the appointment systems literature. The factors included are:

Proportion of planned arrivals. The proportion of planned arrivals was increased in increments of 25% each, starting from 0% (all unappointed arrivals) until 100% (all appointed arrivals). The researchers experimented with more frequent increments. However, the relatively linear relationship between the proportion of appointed arrivals and truck turnaround times meant that more frequent increments would contribute with little additional information.

Arrival punctuality. Punctuality was modelled by adding a stochastic component to each appointed arrival time. The stochastic component consisted of normal distributions with increasing standard deviations, in a similar vein to Ramirez-Nafarrate *et al.* (2017):

(1) High punctuality: 95% of arrivals are within +/-5 min. from appointment time {N(0, 2.5)},

(2) Medium punctuality: 68% of arrivals are within +/-5 min. from appointment time {N(0, 5)},

(3) Low punctuality : 38% of arrivals are within +/-5 min. from the appointment time {N(0,10)},

The scenario analysis included combinations of the 3 factors and resulted in 13 scenarios. Each scenario was run 50 times and each iteration simulated a year of operations.

The appointment system simulated 6 appointments per time-window (the length of a time-window is 1 hour). This generated a similar truck arrival average (inter-arrival time of ~10 minutes/truck) to that observed at the terminal.

RESULTS

The baseline scenario – marked with the black diamond symbol – in which all trucks arrive unappointed at the terminal simulates the empirical situation currently observed. The average truck turnaround time in this scenario is approximately 22.5 minutes and includes, on average, 6 minutes of waiting time per trucks. Most of the waiting time is accrued while truck operators wait for an unloading ramp to become available.

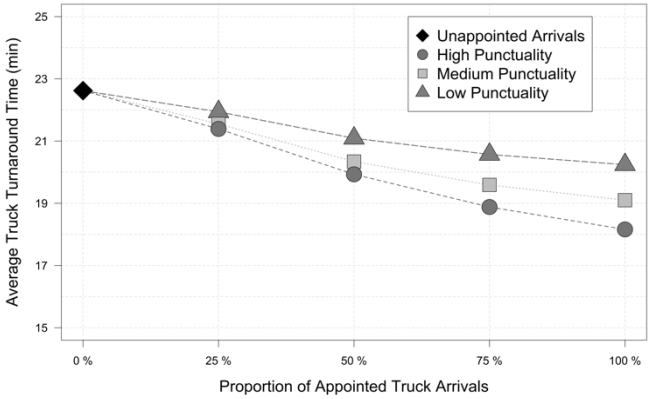


Figure 3: Simulation Model Scenario Analysis Results in Terms of Average Truck Turnaround Times, Truck Arrival Punctuality, and Proportion of Appointed Arrivals

The results of the scenario analysis are presented in Figure 3. With each 25% increase in the proportion of appointed arrivals, truck turnaround times decrease by approximately 6% in the high punctuality scenarios and approximately 3% in the low punctuality scenarios. A 20% reduction in turnaround times is expected when all drivers arrive appointed and with high punctuality. The reduction in turnaround times decreases to 16% with medium punctuality and down to 11% with low appointment punctuality. Turnaround times are primarily improved through reductions in waiting times.

The comparison between high punctuality but lower proportion of appointed arrivals also generates interesting insights. 25% of highly punctual appointed arrivals can generate a slightly lower improvement in turnaround time than 50% of low punctuality appointments. This relation is

reversed in subsequent system use scenarios. 50% of highly punctual appointed arrivals can generate slightly better turnaround times than the subsequent low punctuality scenarios.

DISCUSSION AND FUTURE RESEARCH

Simulation modelling results revealed that both the proportion of appointed arrivals and the arrival punctuality contributed to reducing truck turnaround time. Truck turnaround times gradually decrease as the proportion of appointed arrivals increases, in all arrival punctuality scenarios. However, when only a quarter of arrivals are appointed, the impact of punctuality is relatively small. Therefore, particularly in the early stages of an appointment system implementation, it is important to ensure that the system gains acceptance and is adopted by users.

Arrival punctuality plays an important role in reducing turnaround times as the appointment system adoption increases. Improving the punctuality of 50% of trucks arriving with appointments generates more benefits in terms of truck turnaround times than increasing the proportion of appointed arrivals to 75%.

The results of the simulation model as well as the repeated interaction and feedback cycle with the terminal and its users played an important role in supporting the terminal's decision to start the procurement process for a terminal appointment system.

Future research aims to follow the implementation process of the TAS and its impact. From a modelling perspective, the impact of punctuality and system utilization in a range of throughput scenarios as well as the impact of different truck service rules on truck turnaround times and the wood chip supply chains will also be explored. This research is part of an ongoing project centred on multiple wood chip terminals in Australia on the use of information and digital systems to address land-side congestion at marine terminals. The researchers aim to use the simulation model in other terminals to confirm the results obtained in this case or identify differentiating factors.

CONCLUSION

This research presented a discrete event simulation model of a bulk cargo marine terminal developed based on empirical data from weigh-bridges and truck-mounted GPS units to understand the impact of the proportion of appointed arrivals and arrival punctuality on truck turnaround times.

The scenario analysis results indicate that both the proportion of appointed arrivals and arrival punctuality can reduce truck turnaround times. In the early stages of adoption, increasing the system use can be more beneficial than ensuring punctual arrivals. However, as the proportion of appointed arrivals increases, truck arrival punctuality plays an increasingly important role in reducing service times.

It is important to acknowledge some of the limitations of this research, which primarily related to scope, behaviour modelling and generalizability. The scope of the model is restricted to the gate and unloading operations. Whilst a more comprehensive model was desirable, the data to generate such a model were unavailable. Truck operators display a wide range of behaviours during their daily tasks of which only a small proportion can be geographically captured and simulated. Furthermore, the introduction of a new technology can alter their behaviours. However, these changes in behaviours over time cannot be easily captured and considered in the modelling. Finally, bulk cargo operations are relatively less complex than container operations, on which the majority of the congestion management literature is based. Therefore, one should consider the degree of generalizability of these findings to other environments with care.

This research is part of an ongoing project undertaken in Australia funded by the Australian Research Council through the Industrial Transformation Research Program 'The Centre for Forest Value' https://www.utas.edu.au/arc-forest-value.

REFERENCES

Ackoff, Russell. 1978. *The Art of Problem Solving*. New York, Chichester, Brisbane, Toronto, Singapore: John Wiley & Sons.

Chen, Gang, Kannan Govindan, and Zhongzhen Yang. 2013. "Managing Truck Arrivals with Time Windows to Alleviate Gate Congestion at Container Terminals." *International Journal of Production Economics* 141 (1): 179–88.

Chen, Gang, and Liping Jiang. 2016. "Managing Customer Arrivals with Time Windows: A Case of Truck Arrivals at a Congested Container Terminal." *Annals of Operations Research* 244 (2): 349–65.

Guan, Chang, and Rongfang Liu. 2009. "Modelling Gate Congestion of Marine Container Terminals, Truck Waiting Cost, and Optimization." *Transportation Research Record: Journal of the Transportation Research Board* 2100 (2100): 58–67.

Huynh, Nathan, Daniel Smith, and Frank Harder. 2016. "Truck Appointment Systems." *Transportation Research Record: Journal of the Transportation Research Board* 2548: 1–9.

Huynh, Nathan, and C. Michael Walton. 2008. "Robust Scheduling of Truck Arrivals at Marine Container Terminals." *Journal of Transportation Engineering* 134 (8): 347–53.

Huynh, Nathan, and Michael Walton. 2011. "Improving Efficiency of Drayage Operations at Seaport Container Terminals through the Use of an Appointment System." In *Handbook of Terminal Planning*, 323–44. New York, NY: Springer.

Li, Na, Gang Chen, Kannan Govindan, and Zhihong Jin. 2018. "Disruption Management for Truck Appointment System at a Container Terminal: A Green Initiative." *Transportation Research Part D: Transport and Environment* 61: 261–73.

Morais, P, and E Lord. 2006. "Terminal Appointment System Study." *Transportation Research Board* 1 (March): 123.

Neagoe, Mihai, Mohammad Sadegh Taskhiri, Hong-Oanh Nguyen, Hans-Henrik Hvolby, and Paul Turner. 2018. "Exploring Congestion Impact beyond the Bulk Cargo Terminal Gate." In *Logistics 4.0 and Sustainable Supply Chain Management, Proceedings of HICL 2018*, 63–82.

Ramírez-Nafarrate, Adrián, Rosa G. González-Ramírez, Neale R. Smith, Roberto Guerra-Olivares, and Stefan Voß. 2017. "Impact on Yard Efficiency of a Truck Appointment System for a Port Terminal." *Annals of Operations Research* 258 (2): 195–216.

Torkjazi, Mohammad, Nathan Huynh, and Samaneh Shiri. 2018. "Truck Appointment Systems Considering Impact to Drayage Truck Tours." *Transportation Research Part E: Logistics and Transportation Review* 116 (September 2017): 208–28.

AN APPROXIMATE INFERENCE APPROACH FOR AUTOMATED SUPPLY CHAIN FORMATION

Florina Livia Covaci
Business Information Systems Department
"Babeș-Bolyai "University
Cluj-Napoca,
Romania
E-mail: florina.covaci@ubbcluj.ro

KEYWORDS
Supply Chain, Message Passing, Approximate Inference, Graphical Models

ABSTRACT

The new industrial revolution will transform the traditional supply chain enviroment towards a digital one where autonomous machines will enable the formation of supply chain and will enhance coordination between the participants. The current paper proposes a mechanism based on a message passing inference scheme for solving the digital supply chain formation problem. Within the process of supply chain formation, the trading entites must agree on multiple contract parameter values that must be propagated across the whole supply chain. We create our model based on graphical models theory using intelligent agents that act as participants in the supply chain. We propose an approximate inference scheme for communication between agents along the supply chain in order to obtain a good performace regarding communication requirements.

INTRODUCTION

The new industry revolution will require automated mechanisms to form and coordinate supply chains. The formation of the supply chain in the initial stage determines its future performance, therefore it is demanding to identify mechanisms that will allow the formation and coordination of supply chains autonomously.

There are several approaches in the literature that tackle the Supply Chain Formation (SCF) based on multi-agent systems architectures. In the literature there can be found several contributions where participants are represented by computational agents that act in behalf of the participants during the SCF process.

The first approach that was trying to solve the SCF problem in a fully decentralized manner was the work of (Winsper and Chli 2012). The authors proposed a decentralized inference scheme, named Loopy Belief Propagation (LBP) using a message passing mechanism comparable to the placing of bids in standard auction-based approaches. The paper uses an iterative message passing scheme for the estimation of marginal probabilities of nodes: at each iteration, each node in the graph sends a message to each of its neighbors giving an estimation of the sender's beliefs about the likelihoods of the recipient being in each of its possible states. Nodes then update their beliefs about their own states based upon the content of these messages, and the cycle of message passing and belief update continues until the beliefs of each node become stable.

As LBP suffers of scalability issues in (Penya-Alba, et al., 2012) the authors introduce the Reduced Binary Loopy Belief Propagation algorithm (RB-LBP). RB-LBP is based on the max-sum algorithm and introduces binary variables in order to encode decoupled buy and sell decisions and a selection term and an equality term to assure coherent decisions between participants.

A negotiation-based task allocation method was proposed in (Kong, et al., 2015) in which every agent owns only a local view and the potential resources are found by consumers through peer-to-peer relationships. After the consumers find the potential resources, they start to negotiate with the resource providers. It is often difficult for agents to decide the optimal contract prices, so the agents have the option to negotiate with more than one potential partner, and thus the de-commitment and penalties are necessary and considered in the negotiation process.

In (Kong, et al., 2016) a decentralized approach in which providers and consumers are modelled as intelligent agents was proposed for group task allocation in dynamic environments. The proposed approach allows agents to enter and leave the environments at any time and the tasks have deadlines, and may need the collaboration of a group of self-interested providers. Each consumer has the role of an auctioneer for itself in order gather all the resources required by its task, no central authority/auctioneer being involved.

A pruning-decomposition loopy belief propagation-based method, called PD-LBP, was proposed in (Kong, et al., 2017) for task allocation in dynamic environments. It is composed of two phases: a pruning phase that aims at reducing the searched resource providers, and a decomposition phase that decomposes the initial network into several independent sub-networks on which is operated in parallel the belief propagation algorithm. As opposed to LBP where only the quotes of the participants are considered, the PD-LBP approach considers both a reserve price and a deadline for agreement to be accomplished.

Against this background, the current paper proposes an approximate inference mechanism extending our previous work in (Covaci, 2017a) (Covaci, 2017b) (Covaci & Zarate, 2019) that used a descentralized mechanism and utility functions as a means for incorporating in the supply chain formation mechanism participants preference upon multiple contract parameters (e.g. quality issues, delivery constraints etc.). Although our previous work advanced the state of the art being able to deal with multiple and flexible contract parameters and the assessment of the optimal supply chain

was based on utility functions, it encounters performance issues when contract parameters take values on continuous domains.

In order to overcome this issue we propose the use of approximate messages within a cluster graph propagation mechanism. Instead of representing messages in the cluster graph as factors, we use more compact representations of approximate factors.

This paper is organized as follows. Related work to SCF is presented in Section 2, Section 3 provides the mathematical background needed to describe our approach, Section 4 introduces our approximate inference proposed approach. Evaluation is presented in Section 5, and the conclusions in Section 6.

RELATED WORK TO AUTOMATED SUPPLY CHAIN FORMATION

Several approaches regarding Supply Chain Formation (SCF) problem have been proposed in the multi-agent systems community that are using computational agents on behalf of the participants during the SCF process and automate the SCs formation.

The approaches of (Walsh and Wellman 2000), (Cerquides et.al. 2007), (Giovannucci et. al. 2008), (Mikhaylov et. al. 2011) were proposing to solve the supply chain formation problem in a centralized way. The participant agents informed a central authority of their preferences that were encoded as offers. After collecting the offers of all participant agents, the central authority determined the resulting SC.

In order to overcome the limitations of the centralized approaches mediated SCF approaches were proposed. In this setting, participant agents resort to local markets in which the goods they want to sell or buy are being traded (Walsh and Wellman 2003). The authors proposed a market protocol with referred to as simultaneous ascending (M+1)st price with simple bidding (SAMP-SB), that was using multiple simultaneous ascending double auctions. SAMP-SB was able to produce highly-valued solutions that were maximizing the difference between the costs of participating producers and the values obtained by participating consumers over several network structures. The authors also proposed a similar protocol, SAMP-SB-D, that included a de-commitment phase in order to remedy the solutions in which one or more producers acquire an incomplete set of complementary input goods and were unable to produce their output good, leading to negative utility.

In order to provide better scalability for large scenarios, descentralized peer-to-peer approaches have been proposed.

Loopy Belief Propagation (LBP) is the first peer-to-peer approach that has been used to solve the SCF problem in a decentralized manner (Winsper and Chli 2010), (Winsper and Chli 2012), (Winsper and Chli 2013). The work in (Winsper and Chli 2013) shows that the SCF problem can be cast as an optimization problem that can be efficiently approximated using max-sum algorithm for loopy graphs or can find exact solutions when the graph is a tree. LBP starts by initializing the beliefs of each agent about each of their possible states to zero. Each agent then passes a message containing a vector of belief values to each of its neighbors in the network. Once all agents have passed a message to each of their neighbors, each agent updates its beliefs based upon the content of the messages it received. The cycle of message passing and belief update continues until the network becomes stable when finally, the states of the variables are determined.

In order to overcome scalability issues present in (Winsper and Chli 2010), (Winsper and Chli 2012), in (Penya-Alba, et al., 2012) the authors propose an approach for decentralized SCF, called the Reduced Binarized Loopy Belief Propagation algorithm (RB-LPB). RB-LBP is based on the max-sum algorithm and is able to simplify the computation of max-sum messages, hence reduces the computation required in to assess the supply chains.

In order to improve the problem solving time (Kong, et al., 2017) proposed a pruning-decomposition loopy belief propagation-based method, called PD-LBP. The prosed method included two pahses: a pruning phase for reducing the searched resource providers that do not meet the reserve value, and a decomposition phase for decomposing the initial network into several independent sub-networks on which is operated in parallel the belief propagation algorithm.

MATHEMATICAL BACKGROUND

This section will present the concepts needed to encode the SCF problem within our approach. Hence in the following we will describe the theoretical foundation regarding the probabilistic graphical models that allow pursing inference.

Graphical models are centered around the idea of factorization. As indicated by (Wainwright & Jordan 2008), such a model encodes a series of probability distribution functions which factorize according to a certain configuration of the graph.

In the following we define the factors and major operations of our interest on the factors. For precision, we exactly follow the well-accepted formalization of (Koller & Friedman 2009).

Definition (Factors) (Koller & Friedman 2009) : Let D be a set of random variables. A factor θ is a function from Val(D) to IR. A factor is nonnegative if all its entries are nonnegative. The set of variables D is called the scope the factor, and it is denoted by Scope[θ].

When performing computation of the probability of a subset of variables, the key operation which is performed is the marginalization the variables of a distribution.

Definition (Factor marginalization) (Koller & Friedman 2009): Let X be a set of variables, a variable $Y \notin X$ and a factor $\theta(X; Y)$. The factor marginalization of Y in θ, denoted $\sum_Y \theta$, is defined to be a novel factor over X: $\psi(X) = \sum_Y \theta(X,Y)$.

In other words we can call this operation summing out of Y in ψ.

As observing from the definition above, computing the marginal distribution over a subset of X could be viewed as an operation on a factor.

Next we will we will defne the notion of the cluster graph, as being a data structure suitable for the factor-manipulation process in a graphical way. For each subset of variables in the graph we will associate a node in the cluster graph. Undirected edges will connect the nodes in the cluster graph who have a non-empty intersection of the variables within

their scope(Koller & Friedman 2009). The formal definition is given below:

Definition (Cluster graph) (Koller & Friedman 2009) : A cluster graph U for a set of factors θ over X is an undirected graph. Each node *i* of the cluster graph is a cluster associated with a subset Ci ⊆ X. A cluster graph must be family-preserving, meaning that each factor θ⊆Θ must be associated with a cluster C_i, denoted α(θ), such that Scope[θ] ⊆ C_i. Each edge between a pair of clusters C_i and C_j is associated with a sepset $S_{i,j} \subseteq Ci \cap Cj$.

In situations where exact inference is infeasible there is need to perform perform a projection to approximate a complex distribution with one with a simple structure. In such an approximation we would like to find the best approximation within a family in which we can perform inference.

Suppose we have a distribution P and we want to approximate it with another distribution Q in a class of distributions Q.

Definition (M-projection) (Koller & Friedman 2009) : The M-projection (moment projection) of P onto Q is the distribution Q^M = arg min D(P|Q).

Let P be a distribution over X1,…,Xn, and let Q be the family of distributions consistent with G, the empty graph. Then Q^M = arg min D(P|Q) is the distribution Q^M(X1,…,Xn) = P(X1)P(X2) … P(Xn).

Hence, the M-projection of P onto factored distribution is simply the product of marginals of P.

It can be observed that the family Q of fully factored distributions is characterized by a vector that simply counts, for each variable Xi, the number of occurrences of each of its values. The marginal distributions over the X_i's are simply the expectations, relative to P. We see that, by selecting Q to match these expectations, we obtain the M-projection.

PROPOSED APPROACH

In order to model the supply chain formation process in the current paper, we extend the work presented in our previous papers (Covaci, 2017a) (Covaci, 2017b) (Covaci & Zarate, 2019) by addresing the issue of contract parameters taking values on continuous domains. As pointed in (Covaci & Zarate, 2019) although our approach brought several benefits over the state of the art approaches, it ecounters performance issues when contract parameters take values on continuous domains. The memory needed to store the preferences of the participants and the communication requirements will grow with the size of the parameters' domains involved in the utility functions of the participants in the supply chain.

Hence the present paper proposes an approximate inference approach in order to address the performance issues when contract paramters take values over contiuous domains.

The supply chain is modelled using a graph in which the nodes represent the trading enitites in the supply chain and the edges represent the possible trading relationships between the entitites. The entities might have the role of suppliers, consumers or both roles when the entities are part of the middle tier network.

The entities are acted by intelligent agents that are using a message passing mechanism in the graph in order to agree on contract parameters. The agents talk with each other about multiple contract parameters and they have to agree on a contract that is composed of the actual values of the issues that they have talked about. Each participant has certain preferences upon different contract parameters. The agreed values of the negotiated issues are reflected in a contract which has a certain utility value for every agent. By using utility functions, they can assess the benefits they would gain from a given contract, and compare them with their own expectations in order to make decisions. Moreover during the supply chain formation process, the agents send messages to their neighbors regarding the states of their variables that are maximizing their own utility functions.

In order to find the best possible assignment in the supply chain network, we follow a two steps mechanism: 1. The transformation of the initial graph into a cluster graph; 2: Passing of messages over the cluster graph using max-sum algorithm. The messages are scheduled from leaves to root and back in order to be able to compute messages using two passes in the graph.

For the illustration we present a simplified example from the automotive industry as the one in Figure 1.

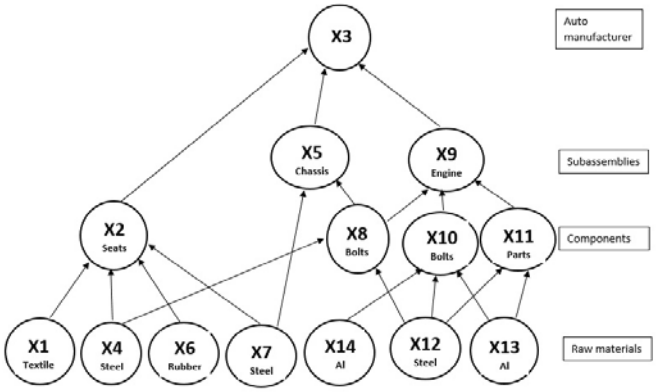


Figure 1 Simplified example graph for automotive supply chain environment

The auto manufacturer needs to procure several components and subassemblies from possible suppliers in order to manufacture the car. The components and subassembly supplier also need to procure raw materials from lower level suplliers. Taking a look for example at X2 (the seats provider) he needs to procure textile, steel and rubber from one ore more underlying suppliers in order to deliver the seats the automanufacturer.

The nodes in the supply chain graph are possible trading partners. It is assumed that they can provide the required products/services as suppliers or they are willing to buy the products/services in their quality of consumers in the supply chain network. Each possible participant in the supply chain owns a utility function in order to model his own preferences upon the contract parameters that he negotiates on. For all the exchanged messages it will be considered that each utility function of the participants depends on multiple contract parameters.

In order to transform the initial graph into a cluster graph will consider the utility functions as factors. Will assign each factor to the corresponding node in the supply chain network and remove the arrows. Each edge in the cluster graph means

that the clusters are sharing one or more variables that they are talking about.
The transformation of the initial graph into a cluster graph is presented in Figure 2.

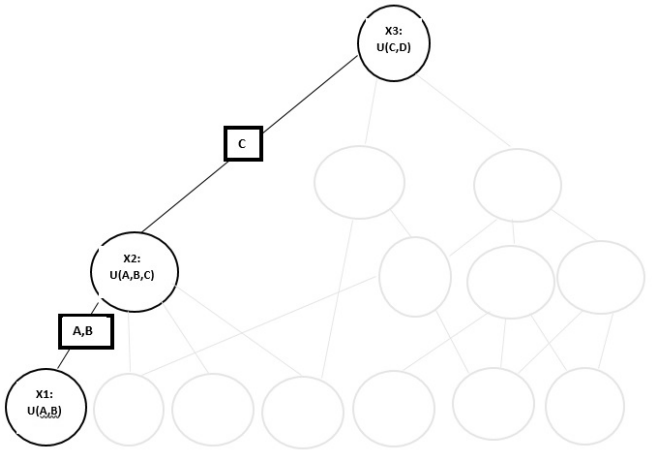


Figure 2 Subgraph of obtained cluster graph

The second step in the proposed mechanism consists of message passing over the cluster graph according to the equations (1) and (2).
By sending the message in equation (2), X_1 says to X_2 which is his preferred value from the set of values for issue B.

$$\lambda_{1\to 2}(B) = max_B(U(b_j, c_k) + max_A\left(U(a_i, b_j)\right)) \quad \textbf{(1)}$$

X_1 sends the max-marginalization of B over A ($max_A(U(a_i,b_j))$ and then adds the computed utility of agent X_2 and then computes the max marginalization of B over the above terms.
Agent at node X_2 evaluates using his utility function, the utility that he gets for each combination of values from the set of values for issues B and C. X_2 send to X_1 the message in equation (2), which is his preferred value from the set of values for issue B.

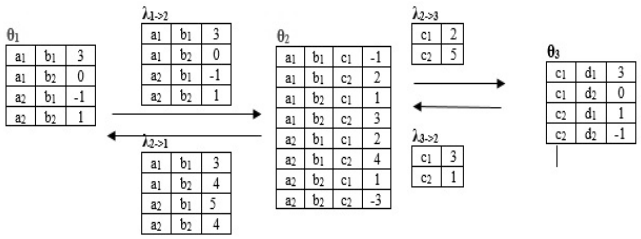


Figure 3a X1-X2-X3 Message exchange

X_2 sends the max-marginalization of B over C (max_C (U(b_j ,c_k)) and then adds the computed value for utility of agent X_1 and then computes the max marginalization of B over the above terms.

$$\lambda_{2\to 1}(B) = max_B(U(a_i, b_j) + max_C\left(U(b_j, c_k)\right) \quad \textbf{(2)}$$

The messages are scheduled starting from leaves and then are propagated upward towards the root.
Figure 2 and 3 illustrate an example of message exchange over cluster graph where each parameter can have only two values.
Following Figures 3a and 4 we have to point out that when an agent has utility function with two parameters and each parameter cand have two values (e.g. X1), the number of possible joint assignments is four.
However as we can observe in Figure 3b, if the agent X1 owns a utility function with two parameters and each parameter take four values, the message sent from X1 to X2 contains sixteen possible joint assignments. As the larger the domain of the variable will be the size of the message will increase. Hence, when dealing with large continuous domains the needed bandwith for communication between the agents will be increase dramatically.

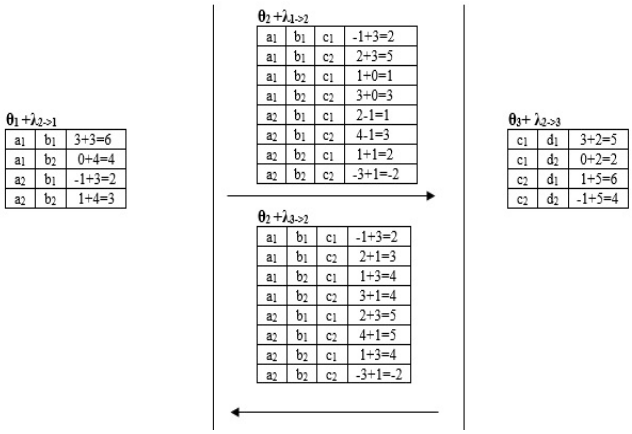


Figure 3b Computation performed during the X1-X2-X3 message exchange

Consequently we need to simplify the messages by using approximate messages in cluster graph propagation rather than exact messages. Instead of representing messages in the clique trees as factors, we use more compact representation of approximate factors. Alternatively of representing the message with values for the sixteen possible joint assignments, we send a message with two values for each variables, leading to a total number of eight values sent. If we take in consideration continuous domains for the contract parameters, then the saving will be more substantial.

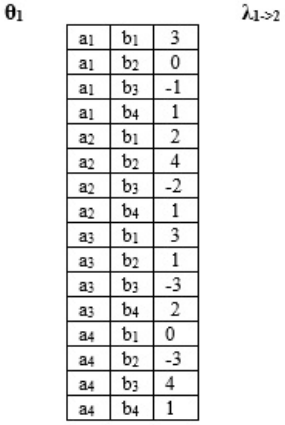


Figure 4 Joint assignments with two paramters and four values for each

As pointed in equations (1) and (2), the proposed approach involves two main operations: summing and max-marginalization. When performing marginalization $max_A(U(a_i,b_i))$ in message $\lambda_{1->2}$, this operation couples all the factors that contain A.
To achieve the effcient inference steps we discussed earlier, we want to approximate the message $\lambda_{1->2}$ *by* a product of

simpler terms.
We introduce the notation M-proj$_{i,j}$ to be the combined operation of marginalizing variables that are not shared by two adjacent agents and performing the M-projection.
Considering the message passing in an exact inference approach, there is need to perform the following operations:

$\lambda_{1->2} = \theta_1$

$\lambda_{2->3} = \max_B (\theta_{2+} \lambda_{1->2})$

In order to reduce the computational complexity we use approximation by replacing each meassage with the M-projected version:

$\lambda'_{1->2} = \text{M-proj-d}_{1,2}(\theta_1)$

$\lambda'_{2->3} = \text{M-proj-d}_{2,3}(\theta_{2+} \lambda'_{1->2})$

Hence, rather than simply marginalizing the sum of factors, we computes their M-projection.
Each cluster collects the messages from its neighbors and sends outgoing messages and this process converges after a single upward-and-downward pass over the cluster graph structure.
However, we must note that the impoved computational complexity comes with a issue. This approximation makes conditional independence assumptions about the variables. Therefore in the example provide in Figure 1, the utility functions of the agent X1, that provides textile for the car seats, incorporates two parameters (A,B). Let A be the colour of the textile and let B the time needed to deliver the textile. When using exact inference as in our previous work we are able to capture the dependencies between the two parameters as the car seat producer X2 may prefer to get brown textile in a shorter time rather than black textile but with a longer delivering time.
When using approximate inference the conditional dependence information is lost as the car seat provider X2 might prefer black textile as it is more likely to be required by the final customers and loosing the information about long delivering time.

EVALUATION AND SIMULATION

Our previous work regarding SCF was using exact inference for the message passing mechanism and encountered performance issues when parameters were taking values over large continuous domains. The current approach overcomes this limitation by proposing in the section above a mechanism for automated SCF using multiple contract parameters that uses approximate inference for sending messages aiming at improving the communication performace. Therefore, in the following we will analyze the communication requirements for sending the messages over the clusters in our cluster graph for the both approaches.
In the proposed approach two agents that are interested in establishing a commercial relationship, exchange a message regarding the variables they share in their utility functions. Hence the communication requirements depend on the number of shared parameters and the number of the values that these paramters can take.

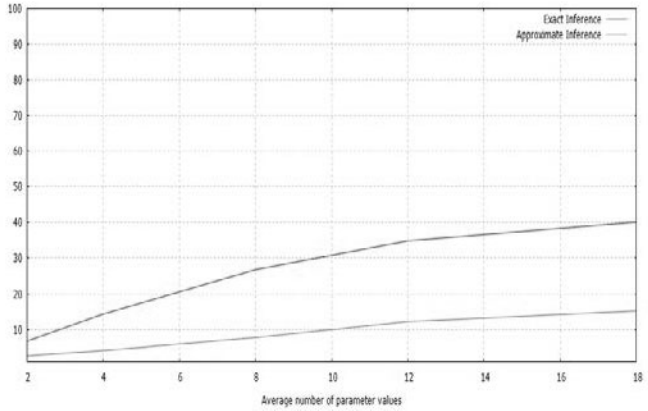


Figure 5 Communication requirements (KB) – (1-2 shared parameters)

For the analysis of the above proposed mechanism we have run simulation on two type of datasets: datasets regarding agents that share 1-2 paramters (Figure 5) and parameters that share 3-4 paramters (Figure 6).

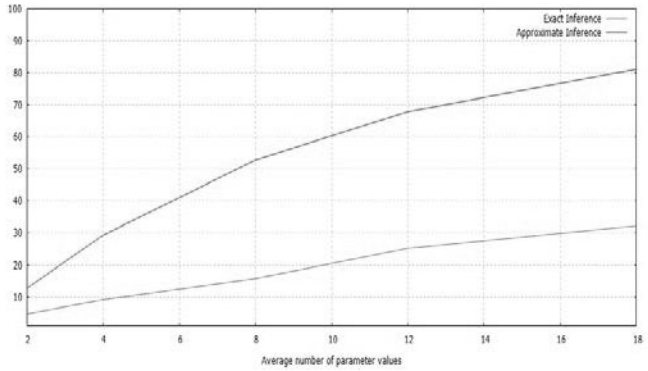


Figure 6 Communication requirements (KB) – (3-4 shared parameters)

The runned simulation show the that communications requirements increase according to the number of shared parameters. Moreover when using an approximate inference approach there is a significant improvement in the performance regarding communication requirements compared to exact inference approach. Howerver, as noted in the above section, this performance improvement comes with the drawback of the independence asspumtion between parameters in a supply chain contract.

CONCLUSIONS AND FUTURE WORK

The present paper proposes an automated mechanism for supply chain formation that employs a message passing mechanism allowing participants to share their beliefs about the optimal contract paramters values of the supply chain only with relevant participants and thus preserving the self-interest of participating agents. Using an approximate inference scheme the proposed approach is able to improve the messages computational complexity and the overall consumed bandwith.

As a future work we consider evaluating various network topologies in diverse economic hypothesis and adding constraints regarding complementary inputs needed for consumer nodes.

REFERENCES

Cerquides, J., Endriss, . U., Giovannucci, . A. & Rodriguez-Aguilar, J. A., 2007. *Bidding languages and winner determination for mixed multi-unit combinatorial auctions.*, IJCAI, Morgan Kaufmann Publishers Inc..

Covaci, F. L., 2017. *A Multi-Agent Negotiation Support System for Supply Chain Formation.* Stuttgart, Hohenheim Discussion Papers in Business, Economics and Social Sciences.

Covaci, F. L., 2017. *Modelling and Simulation for Decentralized Supply Chain Formation.* Lisbon, ESM 2017, Editor: Paulo J.S.Goncalves.

Covaci, F. L. & Zarate, P., 2019. Modelling decision making in digital supply chains: insights from the petroleum industry. *Kybernetes.*

Giovannucci, A., Vinyals, M., Rodriguez-Aguilar, J. & Cerquides, J., 2008. *Computationally-efficient winner determination for mixed multi-unit combinatorial auctions.* s.l., Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems,Volume 2, pages 1071–1078.

Koller, D. & Friedman, N., 2009. *Probabilistic graphical models: principles and techiques.* s.l.:MIT Press.

Kong, Y., Zhang, M. & Ye, D., 2015. A negotiation-based method for task allocation with time constraints in open grid environments. *Concurrency and Computation: Practice & Experience,* 27(3), pp. 735-761.

Kong, Y., Zhang, M. & Ye, D., 2016. An Auction-Based Approach for Group Task Allocation in an Open Network Environment. *The Computer Journal,* 59(3), pp. 403-422.

Kong, . Y., Zhang, M. & Ye, D., 2017. A Belief Propagation-based Method for Task Allocation in Open and Dynamic Cloud Environments. *Knowledge-Based Systems,* Volume 115, pp. 123-132.

Mikhaylov, B., Cerquides, J. & Rodriguez-Aguilar, J., 2011. Solving sequential mixed auctions with integer programming. *Advances in Artificial Intelligence,Springer,* pp. 42-53.

Penya-Alba, T., Vinyals, M., Cerquides, J. & Rodriguez-Aguilar, J., 2012. *A scalable Message-Passing Algorithm for Supply Chain Formation.* s.l., 26th Conference on Artificial Intelligence.

Wainwright, M. & Jordan, M. I., 2008. *Graphical Models, Exponential Families and Variational Inference.* s.l.:Now Publishers Inc..

Walsh , W. E., Wellman, M. P. & Ygge, F., 2000. *Combinatorial auctions for supply chain formation.* s.l., Proceedings of the 2nd ACM conference on Electronic commerce.

Walsh, . W. E. & Wellman, . M. P., 2003. Decentralized supply chain formation: A market protocol and competitive equilibrium analysis. *Journal of Artificial Intelligence Research,* Volume 19.

Winsper, M. & Chli, M., 2010. Decentralised supply chain formation: A belief propagation-based approach. *Agent-Mediated Electronic Commerce.*

Winsper , M. & Chli, . M., 2013. Decentralized supply chain formation using max-sum loopy belief propagation. *Computational Intelligence,* 29(2), pp. 281-309.

The Effects of Storage Assignment and Order Batching Intensities on Picker Blocking in Narrow-Aisle Order Picking Systems

Dirk Kauke and Johannes Fottner
Chair of Materials Handling, Material Flow, Logistics
Technical University Munich (TUM)
Garching,
Germany
email: dirk.kauke@tum.de

KEYWORDS

Order picking systems, narrow aisle, picker blocking, storage assignment, order batching

ABSTRACT

Order Picking Systems account for up to 55 percent of the costs of warehouse logistics in an organisation. The high proportion of manually performed processes especially make a harmonised system necessary. Due to a lack of available space, narrow-aisle order picking systems become more and more relevant. Since the order pickers are not able to pass each other in the aisles, several blocking scenarios can occur. In the past, several studies examined the impact of different order picking planning problems on the intensity of picker blocking. One or two specific planning problems were investigated in most studies. This paper, on the other hand, demonstrates the correlation between different storage assignment strategies and the effect of different intensities of order batching strategies. We will show the required time for order fulfilment as well as the ratio of blocking time for given boundary conditions such as the number of orders and the layout.

INTRODUCTION

Order Picking Systems (OPS) form the centerpiece of modern warehouse operations. Due to the high growth of e-commerce business, OPS have become more and more important. On the other hand, order picking is still very labor- and time-intensive (Frazelle 2010, Tompkins et al. 2010). Experts estimate that almost 80 percent of all warehouses are operated manually and that 55 percent of the total costs can be assigned to the order picking process (de Koster et al. 2007, Napolitano 2012). Due to cost-oriented operations, warehouses are forced to work with the given space even when storage inventory needs to expand. Since the height of shelves in manually operated OPS is limited, one needs to increase space utilization other ways. A reduction of the widths of the aisles to so called narrow-aisles appears to be one common solution (Gue et al. 2006).
However, OPS with narrow-aisles have disadvantages. The individual order pickers cannot pass each other in the aisles so that congestions can occur. The literature distinguishes two different types of picker blocking (Parikh and Meller 2009). In-the-aisle-blocking develops when order pickers cannot pass each other in the aisles. Pick-column blocking describes the scenario when two order pickers want to pick at the same place, which can even occur in wide-aisle OPS (Franzke et al. 2017).
No matter what type of picker blocking occurs, former studies have shown that picker blocking leads to performance losses (Gue et al. 2006, Parikh and Meller 2009). The warehouse processes and common order picking planning problems therefore need to be coordinated.

OBJECTIVE AND STRUCTURE

The main object of this paper is to examine the impact on different storage assignment strategies and order batching on picker blocking in narrow-aisle OPS. We will examine the impact of total order fulfilment time as well as average blocking time per throughput.
The following chapter will provide a brief overview about the literature that has addressed the picker-blocking problem. In the fourth chapter, the relevant influencing parameters that are examined in the discrete simulation model will be introduced. The fifth chapter will present the main results of the simulation. In conclusion, the last chapter will summarize our insights and present an outlook for further studies.

RELATED LITERATURE

Franzke et al. (2017) recently published an extensive literature review about picker blocking in manual order picking systems (Franzke et al. 2017).
Parikh and Meller (2009) developed an analytical model to examine the mentioned pick-column blocking. They showed that, picker blocking in wide-aisle OPS is less

crucial than blocking in narrow-aisle systems (Parikh and Meller 2009).
Elbert et al. (2015) used an agent-based simulation model to investigate the best combination of routing strategies in a rectangular OPS with three order pickers. They showed that, the best routing combination is when two pickers are using the largest gap and one a combined solution (Ebert et al. 2015).
Whereas Franzke et al. (2017) also demonstrated with an agent-based simulation model that the shortest throughput time can be realized when all order pickers follow the same routing policy. They suggest the optimal routing strategy with either the vertical ABC storage strategy or a random storage assignment strategy (Franzke et al. 2017).
Van Gils et al. (2018) investigated picker blocking in a high-level narrow-aisle OPS. In addition to typical horizontal storage assignment strategies, they focused on different vertical storage assignment strategies in a real-life spare parts warehouse. They showed that, next to the currently applied strategies, the order picking process can be performed 6.2 percent more efficiently on average if routing and storage assignment strategies are reconsidered (van Gils et al. 2018b).
Hence, simulation is frequently used to investigate picker blocking in narrow-aisle OPS. No consideration of a combination between storage assignment strategies and batching intensities such as the number of aisles to enter and the stops per aisle has been examined yet. There has also been no detailed investigation of the percentage of blocking time when storage assignment and batching intensities are modified.

INFLUENCING PARAMETERS

Van Gills et.al. published an extensive listing of the major order picking planning problems (van Gils et al. 2018a). Since consideration of all planning problems would exceed the scope of this paper, the main factors will be examined. The parameters can be categorized into two groups. The first group includes all parameters that will not be modified. The rest will be reviewed to identify their impact on picker blocking time.

Fixed parameters

The fixed parameters in the simulation model are:

- Layout design
- Handling equipment
- Routing
- Number of orders per day
- Job assignments

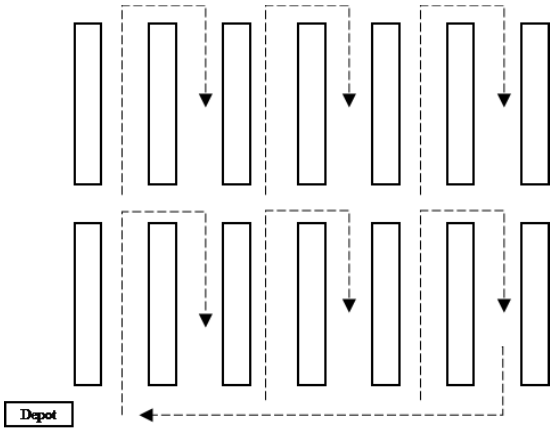


Figure 1: Overview of the examined layout design.

The layout design and the handling equipment are not varied in the simulation runs, because they are mostly fixed at the beginning of the planning phase and will not be changed afterwards (van Gils et al. 2018a).
In many cases the routing of the order pickers is given by the handling equipment and the layout design. In the examined case, the order pickers are using specific order picking trolleys that only allow entry into an aisle one specific way and do not allow to turn within an aisle. Therefore, the order pickers route by the common S-Shape-Strategy with the option to skip aisles (de Koster et al. 2007, van Gils et al. 2018a).

The number of orders per day are fixed at 1,000 to allow comparison between the different strategies. These orders are available at the beginning of every simulation run. No dynamic incoming order process is implemented. The job assignment between order and picker takes place randomly without any prioritization (van Gils et al. 2018a).

Variable parameters

In addition to fixed parameters, there are some others that can be changed between different simulation runs:

- Workforce level
- Storage assignment strategy
- Number of aisles with picking locations
- Stops per aisle

The workforce level describes the available amount of order pickers van Gils et al. 2018a. The different simulation runs will have either five, ten or 15 persons to complete the orders. The considered storage assignment strategies are based on the common strategies that have been examined in several scientific papers (de Koster et al. 2007, Franzke et al. 2017). The figure 2 illustrates the four strategies.

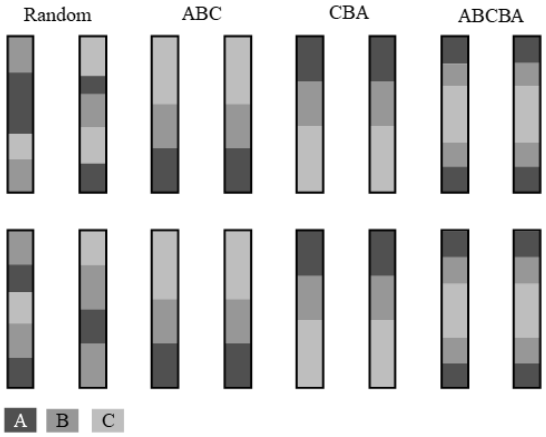


Figure 2: Overview of the examined storage assignment strategies (de Koster et al. 2007, Franzke et al. 2017, Hompel et al. 2011)

The random strategy is the only strategy that allows storage without any restriction. The rest of the strategies consider turnover frequency and combine the articles in specific zones. The ABC strategy stores the high frequency articles (A) at the beginning, the CBA at the end of every aisle. The ABCBA strategy is a mix of the ABC and the CBA where the C-articles are located in the middle of every aisle and the A- and B-articles both at the beginning and the end of every aisle (de Koster et al. 2007, Franzke et al. 2017).
The last two parameters enable examination of the two different batching strategies and finally different long order picking trips. The first one defines the number of aisles the order picker needs to enter to collect the ordered articles. In the studies, the number of aisles is either one or two. The stops per aisle parameter describes how often the order picker is supposed to stop to collect the ordered articles. The smaller the product of number of aisles and stops per aisle is, the more restrictive the actual batching strategy will be to ensure a high utilization of every order picking tour. On the other hand, a batching strategy with more aisles to enter and more stops to pick can result in a higher rate of interactions between the order pickers and possible increased picker blocking. The table 1 summarizes the variable parameters and their values.

Table 1: Summary of examined parameters and their values

Parameter	**Value**
Workforce level	5, 10, 15
Storage assignment strategy	Random, ABC, CBA, ABCBA
Number of aisles with picking locations	1,2
Stops per aisle	2,4,6,8

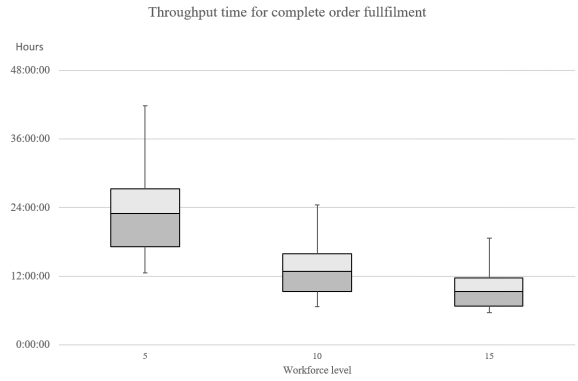


Figure 3: Throughput time with varying parameters

Performance targets

The key figures that identify the performance of the simulation model are:

- Total time to fulfill the orders
- Blocking time per order

The total time to fulfill the orders (throughput time) is a first indicator of how efficient a specific configuration of the order picking system is. Additionally, the ratio between productive and unproductive time is also relevant. The warehouse operator aims to achieve a high ratio of productive time. Therefore, blocking time should be as low as possible (van Gils et al. 2018b).

SIMULATION AND ANALYSIS

The analysis was conducted with the Tecnomatix Plant Simulation software. Plant Simulation is a discrete event simulation tool that is commonly used to investigate material flows and production processes. The simulation model is based on the presented variable and fixed parameters. Every experiment had its defined design set of parameters. To eliminate possible statistical variations, every defined simulation run was replicated five times. The following chapters show the generated results.

General throughput time

The first results should show the development of general throughput time if the variable parameters are changed. Surely an increase of the workforce level should lead to a shorter throughput time. The figure 3 illustrates this relation.
This figure shows that the expected throughput times decrease once the level of available workforces is increased. The single mapped box plots include several storage assignment strategies as well as the number of

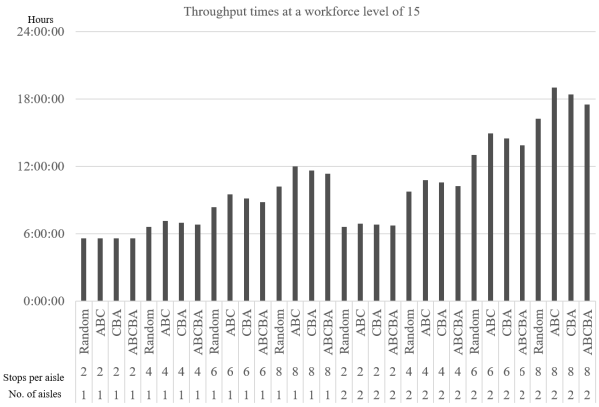


Figure 4: Throughput times at a workforce level of 15

aisles with picking locations as is shown in Table 1. A remarkable result is the wide scattering, especially with 5 people in the order picking system. The longest throughput time of about 42 hours occurs with the ABC storage assignment, two aisles with picking locations and eight stops per aisles. The shortest throughput time of 12.5 hours is achieved by applying a random storage assignment and a two-stop tour in only one aisle.
A workforce level of 15 leads to a significantly narrower scattering. The shortest throughput time is about 5.5 hours and the longest time is about 19 hours. Both the longest and the shortest time can be achieved with the same parameter settings like at a workforce level with five pickers. On closer examination, it should be noted that the throughput time can not be reduced by the same level the workforce level is increased. Adding five more people to the system reduces the average throughput time from 24 hours to 13.5 hours. This becomes even clearer if you look at the development from 10 to 15 people. There, the average time is only reduced to 10.25 hours. This underlines the enormous impact of picker blocking on general performance development.
In summary, the obvious assumption is that a random storage assignment can achieve the shortest throughput time. To support this theory, a deeper look at specific throughput times is needed. Due to the statistical balance and a more likely implementation in practice, the workforce level of 15 shall serve as the basis.

Throughput times at a workforce level of 15

This chapter will provide a more detailed insight in the effects of the storage assignment strategy, number of aisles with picking locations and stops per aisle parameters at a workforce level of 15. The following figure shows the individual throughput times.
This figure illustrates that the shortest throughput time can be achieved if the picker only enters one aisle with two stops. At this point, the differences between the several storage assignment strategies are not clear. All strategies can achieve a throughput time close to six hours. This can be explained by the fact that the given layout has in total 12 aisles with 20 potential picking stops per aisle. Therefore, the residence time per picker and per aisle is very short so that blocking between pickers is less likely. For this reason, a specific storage assignment strategy has a lower impact on the performance.
In addition, if we compare the throughput times of two aisles and two stops with the times of one aisle and two stops, it can be noticed that the expected times are similar. This is also due to the given layout design. Since the single aisles are separated by a crossway, the picker needs to pass trough a lower aisle to get to his/her picking aisle. Thereby it is possible that the picker gets blocked, so that the total time to fulfill the order is close to the time for picking in two different aisles. In this connection, the question must be asked why an obviously more complicated batching strategy with the result of entering only one aisle per tour should be developed, if almost the same performance can be achieved if two aisles with picking locations per tour are allowed.
The diagram 4 also demonstrates that with every defined setting (no. of aisles and stops per aisle) the random storage assignment strategy realizes the shortest throughput. The difference becomes clearer with increasing stops per aisle. On the other hand, the ABC strategy is without exception the worst alternative in regard to expected throughput time. The random strategy allows an equal distribution of all goods so that the picker stops in the aisle at different positions, therefore making a blocking scene less likely.
Also noteworthy is the fact that the CBA set-up enables shorter throughput times than the ABC design. This is because picker blocking is also possible in the middle and top crossways. Once the high frequency A-articles are placed at the beginning in every aisle, a congestion could even develop with pickers that actually want to enter a different aisle.
The ABCBA strategy tries to profit from both ABC and CBA strategies. It can be assumed that this strategy was actually invented as a return or midpoint routing strategy to ensure that the A-articles are located at both entries in every aisle. Since the presented model is based on an S-Shape routing, it has the high density picks located both at the beginning and at the end of every aisle. This still causes a higher rate of picker blocking than the random storage assignment however, with figure 4 showing that it is the second-best strategy.

Development of the throughput time components

Besides high general performance, a warehouse operator tries to achieve balanced utilization (van Gils et al. 2018a). This means firstly, that the capacity of the material handling system, e.g. the trolley, is well equipped

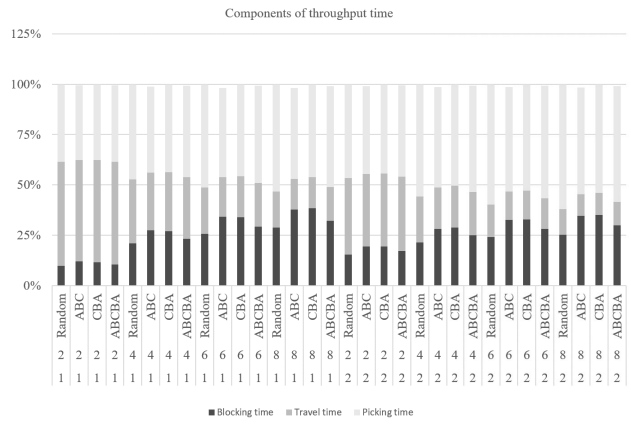


Figure 5: Components of throughput time per order

with orders. Secondly, the order picker performs most of his time on value-adding processes. This means that his total time is spent either preparing the picking order at the depot or picking up goods in the aisles. Whether the travel time between two picks is a value-adding process or not is up to interpretation. On the one hand, the main objective is to reduce travel time in order to increase general performance (Tompkins et al. 2010). On the other hand, in a picker-to-goods-system, the travel time cannot be reduced completely. Additionally, and this fact is especially relevant in narrow-aisle picking systems, if the travel time is reduced by e.g. an aisle with only A-articles close to the depot, picker blocking increases and therefore also total throughput time.
This can be clearly seen in the following figure. This figure illustrates the percentage of the three throughput time components blocking time, travel time and picking time. Figure 4 showes that general throughput time is increased by adding more aisles and stops per aisle. Figure 5 demonstrates how this affects the various components.

The graph of blocking times increases from two stops per aisle to eight stops per aisle in the case of one as well as two aisles with picking locations. The lowest value is approximately ten percent, the highest approximately 40 percent. In both cases, the lowest value can be achieved with a random storage assignment strategy and only two stops per aisle, whereas the highest value is at eight stops per aisle and the ABC storage assignment strategy.
As already shown in figure 4, figure 5 also illustrates that the random storage assignment strategy leads to the lowest percentage of blocking time, whereas the ABC and CBA strategies have the highest percentage.
It becomes apparent that there is a linear relation between the increase of throughput time and the increase of blocking time in a narrow-aisle order picking system. On closer examination, it is also noted that the graph of travel time decreases. This is due to the fact that the actual travel time based on the entrance of one or two aisles is fixed. Additionally, the picking time increases with the increase of stops per aisle. Therefore, the percentage of travel time decreases, whereas the percentage of picking time increases.

CONCLUSION AND OUTLOOK

This paper investigates the effects of several storage assignment strategies as well as different order batching intensities on picker blocking in narrow-aisle OPS. A simulation model showed that the shortest throughput time is realized by a random storage assignment strategy regardless of the number of aisles with picking locations and the number of stops per aisle.
The ABC storage assignment in all cases resulted in the longest throughput times since frequent blocking at the beginning of every aisle even causes congestions back to the crossways.
The percentages of blocking time increased equally to general throughput time. The lowest percentages were achieved with a random storage assignment strategy with one aisle and two stops. The highest percentages occurred with eight stops per aisle. In general however, expansion of performance is limited by picker blocking and the given routing strategy. Figure 3 shows that, despite an increase of workforce level, expected performance does not increase at the same rate. This case testifies the need for coordinated OPS.
Future research shall also have a more detailed look at different batching strategies as well as consideration of real order data to validate the generated model and its implemented processes.

REFERENCES

de Koster R.; Duc T.L.; and Roodbergen K., 2007. *Design and Control of Warehouse Order Picking; A Literature review. European Journal of Operational Research*, , no. 182, 481–501. ISSN 03772217.

Ebert R.; Franzke T.; Glock C.; and Grosse E., 2015. *Agent-based analysis of picker blocking in manual order picking systems: Effects of routing combinations on throughput time.* In *2015 Winter Simulation Conference (WSC)*. IEEE. ISBN 978-1-4673-9743-8, 3937–3948. doi:10.1109/WSC.2015.7408549.

Franzke T.; Grosse E.; Glock C.; and Ebert R., 2017. *An investigation of the effects of storage assignment and picker routing on the occurrence of picker blocking in manual picker-to-parts warehouses. The International Journal of Logistics Management*, 28, no. 3, 841–863. ISSN 0957-4093. doi:10.1108/IJLM-04-2016-0095.

Frazelle E.H., 2010. *Trends in Transporation: Logistics management Series.* Atlanta, GA.

Gue K.R.; Meller R.D.; and Skufca J.D., 2006. *The effects of pick density on order picking areas with narrow aisles. IIE Transactions*, 38, no. 10, 859–868. doi:10.1080/07408170600809341.

Hompel M.T.; Sadowsky V.; and Beck M., 2011. *Kommissionierung: Materialflusssysteme 2 - Planung und Berechnung der Kommissionierung in der Logistik.* VDI-Buch. ISBN 978-3-540-29940-0.

Napolitano M., 2012. *2012 Warehouse/DC Operations Survey: Mixed Signals. Modern Materials Handling*, , no. 51, 48–56.

Parikh P.J. and Meller R., 2009. *Estimating picker blocking in wide-aisle order picking systems. IIE Transactions*, 41, no. 3, 232–246. doi:10.1080/07408170802108518.

Tompkins J.A.; White J.; Bozer Y.; and Tanchoco J., 2010. *Facilities Plannning.* New York, NY.

van Gils T.; ; Ramaekers K.; Caris A.; and de Koster R., 2018a. *Designing efficient order picking systems by combining planning problems: State-of-the-art classification and review. European Journal of Operational Research*, 267, no. 1, 1–15. ISSN 03772217. doi:10.1016/j.ejor.2017.09.002.

van Gils T.; Caris A.; and Ramaekers K., 2018b. *Reducing Picker Blocking in a high-level narrow-aisle order picking system. Proceedings of the 2018 Winter Simulation Conference*, 2953–2965.

INVENTORY MANAGEMENT FACING INTERMITTENT DEMAND VIA THE COMPOUND POISSON DISTRIBUTION

Lotte Verdonck
Katrien Ramaekers
Gerrit K. Janssens

Research Group Logistics, Faculty of Business Economics
Universiteit Hasselt – campus Diepenbeek
Agoralaan 1, B-3590 Diepenbeek, Belgium

KEYWORDS
Inventory management, safety inventory, intermittent demand, compound Poisson distribution

ABSTRACT

Intermittent demand is a type of irregular demand including zero demand during a number of time periods. Both forecasting the demand as well as inventory management for these items are different from and more difficult than for items with a smooth and regular demand. First, the use of a compound Poisson distribution is motivated starting from demand from individual customers. Then it is shown how to handle Poisson distributions compounded with uniform and triangular distributions and demand size distributions for intermittent demand cases. Both exact results as approximations, based on simulation results, are proposed. A numerical example is given to illustrate the theoretical results.

INTRODUCTION AND LITERATURE REVIEW

In inventory management, less attention has been paid to irregular demand. This type of demand is characterised by a high level of variability, but may be also of the intermittent type. Intermittent demand items are characterised by occasional demand arrivals interspersed by time intervals during which no demand occurs. Items with intermittent demand include service or spare parts and high-priced capital goods (like in aviation and in defence systems). It also appears with products nearing the end of their life cycle. Demand that is intermittent is often also 'lumpy', meaning that there is also a great variability among the nonzero values (Willemain et al. 2004). A lumpy demand is characterised by a high level of variability, which can be measured by the coefficient of variation.

Furthermore, appropriate models and techniques to manage those items are not often used in real industrial environments, despite active research carried out on this topic in recent years. Both companies and software builders should be aware that inventory management techniques are different for items with different demand patterns. Categorisation of demand patterns consists of defining groups or categories with the purpose of classifying the items to select the best forecasting method and inventory control policy. However, there is no standard classification, definition, or characterisation of the different demand patterns. In order to situate intermittent demand in a classification, the demand pattern types of Williams (1984) and Eaves and Kingman (2004) are used. The classification is based on the Coefficient of Variation of the Demand during Lead Time (CV_{DDLT}), which is made up of three parts: the variation due to the number of orders (CV_n), the variation due to the order size (CV_x), and the variation due to the lead-time (CV_L). The authors use slightly different terminology, as is shown in Table 1.

Table 1: Demand classes based on variation

CV_{DDLT}			*Demand class*	
CV_n	CV_x	CV_L	*Williams*	*Eaves-Kingman*
Low	Low		Smooth	Smooth
Low	High		Smooth	Irregular
High	Low		Slow-moving	Slow-moving
High	High	Low	Sporadic	Mildly Intermittent
High	High	High	Sporadic	Highly Intermittent

Accurate forecasting of demand is important, but the intermittent nature of demand makes forecasting difficult (e.g. Bruggeman et al. 1980; Tavares and Almeida 1983; Watson 1987). The standard forecasting method for intermittent demand items is considered to be Croston's method (Croston, 1972). This method builds estimates taking into account both demand size and the interval between demand occurrences. Despite the theoretical superiority of such an estimation procedure, empirical evidence suggests modest gains in performance when compared with simpler forecasting techniques (Syntetos and Boylan, 2001). Furthermore, the choice of the forecasting method can have an impact on the inventory management policy that is best used. Ramaekers and Janssens (2006) describe an experiment in which the quantitative parameters of an inventory management policy and forecasting method are described. In 2007, Ramaekers and Janssens described a simulation-optimisation model for intermittent demand using a

Markov chain for the occurrence of zero-demand periods. A study by Babai et al. (2010) leads to the conclusion that Naddor's heuristic is the best one able to minimize the total costs associated with inventory policies. Babiloni et al. (2010) summarised, in a table, the type of probability distributions which are used, in literature, for various demand patterns. Their table is reproduced in Table 2.

Table 2:Probability distributions recommended in literature, based on demand characteristics

Demand characteristics	Probability distribution
CV > 0.5 or Right asymmetry of demand shape	Gamma distribution
CV < 0.5 and applies the central limit theorem	Normal distribution
Strategic but slow-moving items	Poisson distribution or Negative Binomial distribution
Items for which probability of no demand during a single period cannot be neglected	Negative Binomial distribution or Compound Poisson distribution or Compound Bernouilli distribution

In the next section, it is shown why the Compound Poisson distribution is useful in an environment in which intermittent demand appears.

MOTIVATION FOR THE USE OF THE COMPOUND POISSON DISTRIBUTION

The approach adopting the Compound Poisson distribution for total demand during a certain period of time starts from the idea that in that period a number of orders are placed and each order size is a random variable from an identical distribution. The number of orders is assumed to follow a Poisson distribution. The total demand D is written as

$$D = Y_1 + Y_2 + \cdots + Y_N \qquad [1]$$

where Y_1, Y_2,.., the individual order sizes, are mutually independent random variables and independent of N.

Assume the demand for a specific product is known to be generated by one or more of a set of m customers. Let the random variable X_k be the demand of customer k for a given period of time, with $k = 1, 2, .. m$. The total demand in a given period is defined by

$$D = X_1 + X_2 + \cdots + X_m \qquad [2]$$

where X_1, X_2, ..., X_m are mutually independent random variables.

To simplify the presentation, it will be assumed that a customer can place only one order per period of time. This assumption is very realistic as it is assumed that orders come in only sporadically.
One can model the individual customer demand X_k as:

$$X_k = I_k B_k \qquad [3]$$

where I_k is an indicator random variable and B_k is the demand size random variable, given that an order occurs. It can be assumed that I_k and B_k are independent. Let

$$p_k = Pr[I_k = 0] \; and \; q_k = Pr[I_k = 1] = 1 - p_k. \qquad [4]$$

The variable I_k follows a Bernouilli distribution with parameter q_k. instead of writing X_k as a product of two independent random variables, one can alternatively write it as the following sum:

$$X_k = \sum_{i=1}^{I_k} B_i \qquad [5]$$

with the usual convention that the value of an empty sum is zero. This representation will be useful to define the 'collective approximation'. The approximation is made by replacing I_k at the top of the summation operator by a Poisson random variable N_k with the same expected value

$$\lambda_k = E[I_k], k = 1,2, \dots, m \; . \qquad [6]$$

Then

$$X_k^A = \sum_1^{N_k} B_i \backsim Compound\ Poisson\left(\lambda_k, f_{B_k}(.)\right) \qquad [7]$$

where $f_{B_k}(.)$ denotes the probability density function of the random variable B_k.

By a known property of the compound Poisson distribution, the random variable

$$D^A = X_1^A + X_2^A + \cdots + X_m^A \qquad [8]$$

has a compound Poisson distribution with parameters

$$\lambda = \lambda_1 + \lambda_2 + \cdots + \lambda_m \qquad [9]$$

and

$$f(x) = \sum_{i=1}^{m} \frac{\lambda_i}{\lambda} f_{B_i}(x) \qquad [10]$$

This means that D^A can be written as

$$D^A = Y_1 + Y_2 + \cdots + Y_N \qquad [11]$$

where Y_1, Y_2,… are independent and identically distributed according to [10] and independent of N which has a Poisson distribution with parameter λ given by [9].

If the variables $X^A{}_k$ are close in distribution to the original X_k, the distribution of D^A will be close to the distribution of D. This happens when every N_k has almost the same distribution as the corresponding I_k. This is true when the

probability q_k of an order is small. And that is exactly what the Poisson approximation of the Binomial distribution stands for.

POISSON DISTRIBUTIONS COMPOUNDED WITH UNIFORM AND TRIANGULAR DISTRIBUTIONS

Given the limited number of observations, it is hard to determine the probability of the DDLT. At most, an expected value and a variance are available. When it is assumed that the order arrival process is a Poisson process, it still leaves us with choice of a probability distribution for the demand size, before any analytics can start. If nothing is known about the shape of the demand size distribution, the maximum entropy distribution with finite support on the finite interval [*a,b*] is the uniform distribution. If one assumes the demand size distribution has a unique mode, then a triangular distribution with support on the interval [*a,b*] might be useful and, if nothing is known about asymmetry, the symmetric triangular distribution seems a reasonable choice. Therefore, in this section, a closer look is taken at Poisson distributions compounded with uniform and symmetric triangular distributions.

Given an expected value and a variance, if the Poisson process is compounded with a uniform distribution, a wide range of distributions showing different values of skewness and kurtosis measures can be obtained. Sometimes, it is even an unreasonable assumption that the demand obeys a known distribution. In such a case, for example with some agricultural products, an inventory replenishment policy has been proposed on the mean of the distribution only (Chen et al., 2016). In Janssens and Ramaekers (2011) an approach has been developed to obtain the reorder point based on the knowledge of the range, mean and variance of the demand distribution only, which is the same information as required for the use of the normal distribution (as many times used in commercial software).

For the purposes of this study, some central moments of the compound Poisson distribution are required. Let the density function of the uniform distribution with support on [*a,b*] be

$$f(x) = \frac{1}{(b-a)} for\ a \le x \le b, 0\ else \qquad [12]$$

When the mean and variance of the aggregated demand are given as M and V, then the lower and upper bound of the uniform distribution can be expressed in their values as:

$$a = \frac{\frac{2M}{\lambda} - \sqrt{\frac{12V}{\lambda} - \frac{12M^2}{\lambda^2}}}{2} \qquad [13]$$

$$b = \frac{\frac{2M}{\lambda} + \sqrt{\frac{12V}{\lambda} - \frac{12M^2}{\lambda^2}}}{2} \qquad [14]$$

From the knowledge that a and $b \ge 0$ and that $M^2 \le E[X^2]$, a valid lower and upper bound for the value of λ can be determined as:

$$\frac{M^2}{V} \le \lambda \le \frac{4M^2}{3V} \qquad [15]$$

Let the density of a symmetric triangular distribution on a finite interval [*a,a+2l*] and mode located at *a+l* be defined as:

$$\begin{aligned} f(x) &= {}^{(x-a)}/_{l^2}\ if\ a \le x \le a+l \\ &= {}^{2}/_{l} - {}^{(x-a)}/_{l^2}\ if\ a+l < x \le a+2l \\ &= 0\ else \end{aligned} \qquad [16]$$

When the mean and variance of the aggregated demand are given as M and V, then the lower bound a and the distance from the bounds till the mode l of the symmetric triangular distribution can be expressed in their values as:

$$a = \frac{M - \sqrt{6(\lambda V - M^2)}}{\lambda} \qquad [17]$$

$$l = \sqrt{\frac{6}{\lambda^2}(\lambda V - M^2)} \qquad [18]$$

For similar reasons as above, bounds on the value of λ can be determined as:

$$\frac{M^2}{V} \le \lambda \le \frac{7M^2}{6V} \qquad [19]$$

Numerical illustration

Consider the following illustration. The demand during lead time has two characteristics: the mean value is kept constant at 100, while the variance ranges from 2000 till 32000. Tables 3 and 4 shows the basic data for ten experiments. Next to the mean and variance also the lower and upper bounds for the order arrival intensity (λ^L and λ^U) are mentioned. In the last two columns the lower and upper bounds for the probability of zero demand (p_0^L and p_0^U) are shown. Table 3 (resp. Table 4) shows these data for the uniform (resp. triangular) distribution as compounding distribution.

Table 3: Basic data for the design with the uniform distribution

Exp No.	*Mean (μ)*	*Variance (σ^2)*	λ^L	λ^U	p_0^L	p_0^U
1	100	2000	5.0000	6.6667	0.0013	0.0067
2	100	4000	2.5000	3.3333	0.0357	0.0821
3	100	8000	1.2500	1.6667	0.1889	0.2865
4	100	16000	0.6250	0.8333	0.4346	0.5353
5	100	32000	0.3125	0.4167	0.6522	0.7316

Table 4: Basic data for the design with the triangular distribution

Exp. No.	*Mean (μ)*	*Variance (σ^2)*	λ^L	λ^U	p_0^L	p_0^U
6	100	2000	5.0000	5.8333	0.0013	0.0029
7	100	4000	2.5000	2.9167	0.0357	0.0541
8	100	8000	1.2500	1.4583	0.1889	0.2326
9	100	16000	0.6250	0.7292	0.4346	0.4823
10	100	32000	0.3125	0.3646	0.6522	0.6945

ORDER-UP-TO-LEVEL POLICY WITH BACKORDERING

In this example the development in Babai et al. (2011) is followed. Inventory is controlled according to a continuous-time order-up-to-level policy. Each replenishment is associated with a deterministic lead-time and unfilled demands are backordered. A penalty cost is incurred if the demand is backordered and a holding cost is incurred if an item is in inventory, both per unit of time.

In the following paragraphs, these notations are used:

λ = demand arrival rate
X = demand size
F = cumulative probability distribution of the random variable X
X_i = the sum of i i.i.d. random variables X ($i \geq 1$)
F^{i*} = cumulative probability distribution of the random variable X_i ($i \geq 1$)
L = lead-time
S = order-up-to-level
h = inventory holding cost per time unit
b = inventory backordering cost per time unit

The probability of i arrivals during the lead-time is

$$p_i = \frac{(\lambda L)^i e^{-\lambda L}}{i!}, for\ all\ i \geq 0 \quad [20]$$

The expected value the total cost, including holding and backordering costs, as a function of S can be written as:

$$E[C(S)] = hSp_0 + \sum_{i=1}^{\infty} \left[h \int_0^S (S-x)\, F^{i*}(x)dx + b \int_0^S (x-S) F^{i*}(x)dx \right] . p_i \quad [21]$$

The objective is to find the optimal order-up-to-level that minimises the expected total cost. Babai et al. (2011) show that the optimal order-up-to-level S is the solution of the equation:

$$\sum_{i=1}^{\infty} \left(\frac{(\lambda L)^i e^{-\lambda L}}{i!} \right) F^{i*}(S) = \frac{b}{h+b} - e^{-\lambda L} \quad [22]$$

Remark that the left-hand side contains an infinite sum, so truncation is required to solve it numerically. Furthermore the complexity of solving this equation lies in the determination of the i-fold convolution of the demand size distribution, F_i. It becomes easier when the distribution F is a type of which all convolutions have the same type of distribution, be it with different parameters. Examples of this type of distributions are the Normal and the Gamma distributions. But neither the uniform nor the triangular distributions fall into this class of distributions.
For the uniform distributions the following is known. The sum of two independent random variables, each having the same uniform distribution, has a symmetric triangular distribution. The density function of the sum of n mutually independent random variables with uniform distribution on [-1,+1] has been determined by Rényi (1970). From this, Kan et al. (2010) determine the distribution on the interval $[a,b]$ as

$$F^{n*}(x) = \sum_{j=0}^{\tau} (-1)^j \frac{1}{j!(n-j)!} \left(\frac{x-na}{b-a} - j \right)^n I_{(na,nb)}(x) + I_{[nb,\infty)}(x) \quad [23]$$

where τ is a function of n, x, a and b. It is the largest integer less than or equal to $(x-na)/(b-a)$.
According to our knowledge, such an expression does not exist for the triangular distribution, except for the fact that the sum of a number of i.i.d. symmetric triangular distributions has also a symmetric distribution. A small simulation experiment with Arena's Input Analyzer has indicated what type of distribution would best fit the simulated sums. The results are shown in Table 5. Sums of two up to five uniform (resp. triangular) distributed random variables have been analysed and some distribution types fitted. In case the table mentions more than one type of distribution, this means that their results are hardly different. The sequence of the multiple types is based on the sequence of the 'square error' measure as shown by the Input Analyzer, from lowest to highest value.

Table 5: Approximate distribution for the n-fold convolutions of uniform and triangular distributions

Number in sum	*Uniform*	*Triangular*
2	Triangular	Beta, Weibull, Normal
3	Normal, Weibull, Beta	Normal, Weibull, Beta
4	Normal, Weibull, Beta	Normal, Weibull, Beta
5	Weibull, Normal, Beta	Normal, Weibull, Beta

Numerical illustration

A situation is considered in which the distribution of an individual demand follows a uniform distribution defined on a finite interval *[a,b]*. The total demand during lead-time has a mean of 100 units and a variance of 8000 (experiment 3 in Table 3). Table 5 shows a set of combinations of λ and *[a,b]* which are valid for the example (from equations [13] to [15]). In Table 6 the optimal order-up-to-level S^* is looked for with cost coefficients $h = 1$ and $b = 10$, $b = 5$ and $b = 1$ by means of equation [22]. The lead-time $L = 1$. F^{2*} is the triangular distribution. For the distributions F^{i*} with $i > 2$, the Normal approximation is chosen (see Table 5). The infinite sum in [22] is cut off at $i = 10$, which offers a precision of 10^{-12}. Note that this approach does not work for point 0 in Table 5 as this compound distribution is a discrete distribution and equation [22] mostly has no solution. In such a case the optimal value S^* might be obtained by the value of S which satisfies the following two conditions

$$\sum_{i=1}^{\infty}\left(\frac{(\lambda L)^i e^{-\lambda L}}{i!}\right)F^{i*}(S) \geq \frac{b}{h+b} - e^{-\lambda L} \qquad [24]$$

and

$$\sum_{i=1}^{\infty}\left(\frac{(\lambda L)^i e^{-\lambda L}}{i!}\right)F^{i*}(S-1) < \frac{b}{h+b} - e^{-\lambda L}$$

Table 6: Valid parameter values for the example based on regular intervals for the parameter λ

Point	*λ*	*a*	*b*
0	1.2500	80.00	80.00
1	1.2917	52.94	101.90
2	1.3333	41.46	108.54
3	1.3750	32.89	112.56
4	1.4167	25.94	115.23
5	1.4583	20.08	117.06
6	1.5000	15.03	118.31
7	1.5417	10.60	119.13
8	1.5833	6.67	119.65
9	1.6250	3.16	119.92
10	1.6667	0.00	120.00

Table 7: Optimal values for S for different combinations of the cost parameters h and b

Point	S^* *(h=1,b=10)*	S^* *(h=1,b=5)*	S^* *(h=1,b=1)*
0	240.0	160.0	80.0
1	232.6	179.9	84.0
2	230.2	184.4	86.5
3	228.7	185.7	87.5
4	227.8	186.2	87.7
5	227.7	186.4	87.7
6	227.6	186.4	87.7
7	227.6	186.4	87.8
8	227.6	186.4	87.8
9	227.6	186.3	87.9
10	227.6	186.2	87.9

CONCLUSIONS

An alternative way of inventory management for intermittent items is proposed. The zero-demand periods are included in the demand distribution during lead-time. The uncertain demand distribution is modeled as a compound Poisson distribution. The number of orders follows a Poisson distribution including zero values. The demand size distribution can be any type of distribution. Here the uniform and triangular distributions are taken as examples. The inventory policy of order-up-to-level policy with backordering is worked out in detail.

REFERENCES

Babai M.Z., Syntetos A.A. & Teunter R. (2010) On the empirical performance of (T,s,S) heuristics, *European Journal of Operational Research,* 202, 466-472.

Babai M.Z., Jemal Z. & Dallery Y. (2011) Analysis of order-up-to-level inventory systems with compound Poisson demand. *European Journal of Operational Research,* 210, 552-558.

Babiloni E., Cardos M., Albarraci J.M. & Palmer M.E. (2010) Demand categorisation, forecasting, and inventory control for intermittent demand items. *South African Journal of Industrial Engineering,* 21(2), 115-130.

Bruggeman W., Muller H. & De Samblanckx, S. (1980) Evaluation of inventory control systems in practice. *Tijdschrift voor Economie en Management* (Leuven, Belgium), 25(1), 19-39.

Chen W., Li J. & Jin X. (2016) The replenishment policy of agri-products with stochastic demand in integrated agricultural supply chains. *Expert Systems with Applications*, 48, 55-66

Croston J.D. (1972) Forecasting and stock control for intermittent demands. *Operational Research Quarterly,* 23, 289-303.

Eaves A.H.C. & Kingsman B.G. (2004) Forecasting the ordering and stock-holding of spare parts. *Journal of the Operational Research Society,* 55(4), 431-437.

Janssens, G.K, & Ramaekers, K. (2011). A linear programming formulation for an inventory management decision problem with a service constraint. *Expert Systems with Applications*, 38, 7929-7934.

Kang J.S., Kim S.L., Kim Y.H. & Jang Y.S. (2010) Generalized convolution of uniform distributions, *Journal of Applied Mathematics & Informatics* (Korea), 28(5-6), 1573-1581.

Ramaekers K. & Janssens G.K. (2006) Towards a best strategy in inventory decision making and demand forecasting for

intermittent demand. In: A. Genco, A. Gentile and S. Sorce (eds.), *Proceedings of the Industrial Simulation Conference*, Palermo, Italy, June 5-7, 2006, pp. 409-414.

Ramaekers K. & Janssens G.K. (2007) A simulation-optimisation framework for determining a best strategy in inventory decision-making and demand forecasting for intermittent demand, *Proceedings of the 2007 European Simulation and Modelling Conference (ESM'2007)*, St. Julian's, Malta, 22-24 October 2007, pp. 249-254.

Rényi A. (1970) *Probability Theory*, North-Holland, Amsterdam.

Syntetos A.A. & Boylan J.E. (2001) On the bias of intermittent demand estimates. *International Journal of Production Economics,* 71, 457-466.

Tavares L.V. & Almeida, L.T. (1983) A binary decision model for the stock control of very slow moving items. *Journal of the Operational Research Society,* 34, 249-252.

Watson R.B. (1987) The effects of demand-forecast fluctuations on customer service and inventory cost when demand is lumpy. *Journal of the Operational Research Society,* 38, 75-82.

Willemain T.R, Smart C.N. & Schwarz H.F. (2004). A new approach to forecasting intermittent demand for service parts inventories. *International Journal of Forecasting*, 20(3), 375-387.

Williams T.M. (1984) Stock control with sporadic and slow-moving demand. *Journal of the Operational Research Society,* 35(10), 939-948.

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX